H. Meschkowski

Unsolved *and* Unsolvable Problems *in* Geometry

UNSOLVED AND UNSOLVABLE PROBLEMS IN GEOMETRY

UNSOLVED AND UNSOLVABLE PROBLEMS IN GEOMETRY

HERBERT MESCHKOWSKI, Ph.D.

Professor of Mathematics, Pädagogische Hochschule, Berlin
Lecturer at the Free University, Berlin

Translated by
JANE A. C. BURLAK

FREDERICK UNGAR PUBLISHING CO.
NEW YORK

American edition 1966
by arrangement with Oliver and Boyd Ltd

First English Edition 1966

© 1966, Oliver & Boyd Ltd

Library of Congress Catalog Card No. 65-25028

PRINTED IN GREAT BRITAIN

PREFACE

At a time when new theories are constantly being developed in mathematics, it seems worth while pointing out that there are still unsolved problems in elementary geometry. We are thinking especially of some problems which are so simple that even an interested layman can understand them. Of course, it is only the formulation of these problems that is quite so simple; and one would need to have attained about the same standard in mathematics as a student in his second year at university to understand the mathematical arguments involved in discussing these problems.

We shall discuss some famous problems which occupied the attention of mathematicians for thousands of years and which have been shown to be unsolvable in recent decades and, in addition, some very old classical problems which, up till now, have scarcely been mentioned in books on the subject. We have endeavoured to vary well-known methods of proof for classical problems and to generalise their formulation.

Obviously a programme like this cannot be fully realised and we have included only those open questions which seemed particularly appropriate. We have completely omitted the problem of " squaring the circle " in its usual form as there are already many descriptions of it. The classical problem of the parallel postulate has also been omitted since the proof that it is unsolvable would require the fundamentals of non-Euclidean geometry to be developed and this would be outside the scope of this book.

In the first chapter we list the problems which we consider in detail. From the consideration of these follow further problems which are formulated later in the text.

Thus, the object of this book is not the systematic representation of a well-defined section of geometry. But rather our aim is of a methodological nature: we want to develop the kind of arguments which are suitable for the solution of the geometrical problems considered and to discuss the questions implied by the existence of unsolvable problems.

The numbers in square brackets refer to the bibliography at the end of the book. This list may be useful for further study. In connection with this we should like to make an important suggestion to the reader.

In the experience of mathematical institutes and authors of mathematical publications attempts are still made to solve problems which are known to be unsolvable. And so we fear that there will be some readers who, after reading this book, will believe that, despite everything, they have found a solution to the problem of trisecting an angle. We should like to advise these readers as follows: Read through the proof showing that such a solution is impossible three times and then try to find the mistake in your argument. Better still, give up such attempts since they really cannot lead anywhere. There are plenty of open questions—many of which are stated in this book—to which mathematical intuition may be applied with real prospect of success. However, it should be noted that it is not particularly easy to solve the open questions stated. It is sometimes much simpler to discover new results in a new branch of mathematics rather than to solve one of the problems left open in elementary geometry. However, according to ERHARD SCHMIDT, it is better to solve old problems with new methods than to solve new problems with old. Therefore we feel justified in encouraging work on the " elementary " problems stated in this book.

For the reader's convenience there is, at the end of the Table of Contents, a diagram showing the interrelationships of the chapters.

HERBERT MESCHKOWSKI

CONTENTS

PREFACE v

Chapter

I. THE PROBLEMS 1

II. REGULAR PACKINGS IN THE PLANE AND ON THE SPHERE 5
- 1. Problems on the density of packing 5
- 2. Lemmas 8
- 3. Packing problems in the plane 13
- 4. Packing problems on the sphere 18
- 5. Further problems 25

III. IRREGULAR PACKINGS AND COVERINGS 29
- 1. The existence theorem 29
- 2. The packing problem graph 30
- 3. A point system for $n = 7$ 32
- 4. The packing problem for $n = 7$ 34
- 5. The covering problem graph 38
- 6. The covering problem for $n = 5$ 40
- 7. The covering problem for $n = 7$ 45
- 8. An outline of further problems 47

IV. PACKING OF CONGRUENT SPHERES 50
- 1. The densest lattice packing 50
- 2. The finest rigid packing 54

V. LEBESGUE'S "TILE" PROBLEM 57
- 1. Formulation of the problem 57
- 2. The hexagonal tile 58
- 3. The tiles of Pál and Sprague 60
- 4. Curves and domains of constant width 62
- 5. Applications 65
- 6. The covering problem in $\mathscr{R}_3$ 67

VI. ON THE EQUIVALENCE OF POLYHEDRA BY DISSECTION 71
- 1. The plane problem 71
- 2. Dehn's theorem 73
- 3. Examples on Dehn's theorem 75
- 4. The algebra of polyhedra 82
- 5. The basis polyhedra 84

VII. THE DECOMPOSITION OF RECTANGLES INTO INCONGRUENT SQUARES 91
- 1. The problem 91
- 2. The definition of a graph 95
- 3. Some calculations for decompositions 97
- 4. The interpretation in electrical engineering 100
- 5. The decomposition of rectangles with commensurable sides 101

VIII. UNSOLVABLE EXTREMAL PROBLEMS 103
1. A theorem on the triangle 103
2. Besicovitch's theorem 107
3. Sets of points at an integral distance apart 109

IX. CONSTRUCTIONS WITH RULER AND COMPASSES 113
1. Some comments on the use of ruler and compasses 113
2. Some construction problems 117
3. The proof of the impossibility of (8a) and (8b) 119
4. The quinquesection of the angle 123
5. The uniqueness of the construction 126

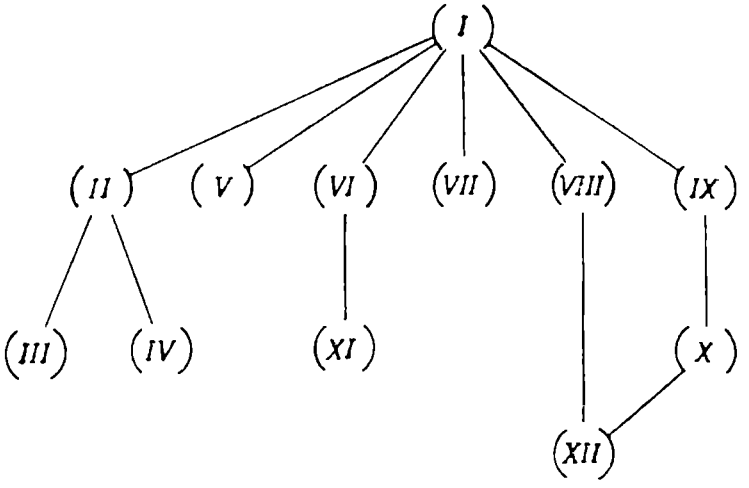
X. CONSTRUCTIONS ON THE SPHERE 128
1. Stereographic projection 128
2. Constructions with spherical compasses and spherical ruler 130
3. Squaring the circle on a sphere 133
4. A lemma 135

XI. PROBLEMS IN THE GEOMETRY OF SETS 137
1. A new definition of equivalence by dissection 137
2. Hausdorff's theorem 141
3. Paradoxical decompositions in $\mathscr{R}_3$ 150
4. Paradoxical decompositions in the plane 155

XII. THE METHODOLOGICAL SIGNIFICANCE OF "UNSOLVABLE" PROBLEMS 159

BIBLIOGRAPHY 163

INDEX 167



Interrelationships of the chapters

Chapter I

THE PROBLEMS

First of all we state the most important of the problems discussed in subsequent chapters. Since this book deals with unsolved or unsolvable problems our main concern is to present the material associated with the solution of these problems. Thus we discuss the solution of related (and usually simpler) problems, give an account of partial results and so on. In the case of unsolvable problems we have *to prove* that no solution is possible. However, we do not indicate in this chapter which of the following problems are unsolvable. We conclude by attempting to explain the significance of the work on unsolvable problems for knowledge and culture in general.

We start then by enumerating the problems. These are numbered for ease of reference. We denote related problems which are introduced later by adding the letters **a**, **b**, ... to the number.

(1) *We are given a spherical planet and* 10 (*n, in general*) *mutually hostile dictators. How can the capital cities of these gentlemen be distributed on the sphere so that the minimum distance between any two of them is as great as possible?*

This problem is equivalent to the following:—What is the smallest radius r of a sphere on which 10 (n, in general) points can be placed at a minimum distance of 1 unit apart?

To see this equivalence it is best to choose as unit of length the radius of the planet. Let δ be the minimum distance apart for any distribution of 10 (n, in general) points on this sphere; then a dilatation by a factor δ^{-1} leads to a sphere of radius $r = \delta^{-1}$ and distance of 1 unit between the points. Thus a minimum value of r corresponds to a maximum value of δ.

This problem (in the second form) was first tackled a few years ago by SCHÜTTE and VAN DER WAERDEN [III 1]. They gave solutions for particular values of n.

(2) *On any spherical planet n* ($n = 9, 10, 11, \ldots$) *fuel depots may be placed so that the distance of any point on the surface from the nearest depot is as small as possible.*

This problem is certainly related to problem (1). It will be seen, however, that the solution of (1) (for a definite value of n) does not by any means always furnish a solution of (2).

(3) *How should one pack a set of congruent spheres (in three-dimensional space) so that the density of packing is as great as possible?*

The " density of packing " can be defined as follows:—If r is the radius of the congruent spheres to be packed without overlapping each other and n_R the number of the spheres of the given set which lie completely within a sphere of radius R, and centre at some origin, then

$$d = \lim_{R \to \infty} \frac{n_R \frac{4}{3}\pi r^3}{\frac{4}{3}\pi R^3} = \lim_{R \to \infty} \frac{n_R r^3}{R^3}$$

is called the density of packing of the set. It can easily be shown (see pp. 7 and 8) that this limit (if it exists) is independent of the particular choice of the origin.

(4) *What is the smallest plane figure with which every cluster of points of diameter* 1 *can be covered?*

A " cluster of points " is a finite set of points and its " diameter " is the greatest distance between any two points of the set.

It is comparatively simple to determine the smallest *circle*, which covers every cluster of points of diameter 1. Its radius, as H. W. E. JUNG [V 2] first showed, is equal to $\frac{1}{3}\sqrt{3}$.

There are, however, still smaller plane figures (we shall call them " tiles ") which cover the finite set of points under consideration. The question, " What is the tile of smallest area? " has not yet been answered.

(5) *Under what hypotheses are polyhedra of equal content also equivalent by dissection?*

Two polyhedra are said to be " equivalent by dissection " if it is possible to decompose them into finitely many sub-polyhedra which

are congruent in pairs and which have no interior points in common:

$$\mathscr{P} = \mathscr{p}_1 + \mathscr{p}_2 + \ldots + \mathscr{p}_n, \quad \mathscr{Q} = \mathscr{q}_1 + \mathscr{q}_2 + \ldots + \mathscr{q}_n, \tag{1}$$

$$\mathscr{p}_\nu \equiv \mathscr{q}_\nu, \quad \nu = 1, 2, \ldots, n.$$

It can easily be shown that plane *polygons* of equal area are also equivalent by dissection.† The corresponding statement for *polyhedra* is not valid. *Necessary* conditions are known for the equivalence by dissection of polyhedra and these are not true for many pairs of polyhedra of equal volume. The problem of finding *sufficient* conditions for equivalence by dissection is not completely solved.

(6) *What is the smallest number of incongruent squares into which a given square can be decomposed?*

It is known that a square can be decomposed into 24 sub-squares, no two of which are congruent. On decomposing a rectangle (with commensurable sides) into squares, no two of which are congruent, it is known that there are at least 9 such squares.

(7) *In a domain $\mathscr{D}$ a line AB of length* 1 *may be moved (by translations and rotations) so that the angle which AB makes with a fixed line is changed by* 360°. *It is required to find a domain of smallest surface area in which this is possible.*

If AB is turned about A through 360° then the line traces out a circle of area π. If AB is restricted to move as a chord of a circle of radius R then the line traces out an annulus with radii R and $\sqrt{R^2 - \frac{1}{4}}$. The area of this annulus is independent of R and equal to $\pi/4$. Next one asks whether there are such domains of still smaller area and which of these domains has minimum surface area.

(8) *Using ruler and compasses a given angle may be divided into five equal parts.*

(9) " *Squaring the circle* " *on the sphere:—Using a spherical ruler and spherical compasses one can construct on the unit sphere a regular quadrilateral which is equal in area to a given spherical cap.*

A " spherical ruler " is an instrument with which one can draw on the sphere the great circle through two given points which are not diametrically opposite. How one actually constructs the circle

† The equivalence by dissection of polygons $\mathscr{P}$ and $\mathscr{Q}$ is defined analogously.

is irrelevant to our problem. It is sufficient to agree that, corresponding to the rules for constructions with ruler and compasses in the plane, it is possible to draw a circle about a (given or already constructed) point P with radius PQ (Q is another given or already constructed point) and that one can draw the great circle through two (not diametrically opposite) points R and S.

Note, however, that both permissible steps of the construction can be carried out using ordinary compasses. To draw the great circle through R and S it is certainly necessary that one should be given the span of the compasses for the great circle ($\sqrt{2}$ on the unit sphere). With the compasses opened to a length of $\sqrt{2}$ one is able first to determine the " pole " T of the great circle RS and then draw about T, again using the span $\sqrt{2}$, the great circle through R and S.

(10) *The set-theoretical squaring problem*:—*The circular disk* $\mathscr{C}$ $(x^2+y^2\leqq 1)$ *and the square* $\mathscr{S}$ *of equal area* $(0\leqq x\leqq\sqrt{\pi};\ 0\leqq y\leqq\sqrt{\pi})$ *may be sub-divided into a finite number n of subsets* $\mathscr{C}_\nu$ *and* $\mathscr{S}_\nu$ *in such a way that the subsets with the same subscript are congruent*:—

$$\mathscr{C} = \mathscr{C}_1+\mathscr{C}_2+\ldots+\mathscr{C}_n,\quad \mathscr{S} = \mathscr{S}_1+\mathscr{S}_2+\ldots+\mathscr{S}_n,$$

$$\mathscr{C}_\nu\equiv\mathscr{S}_\nu,\quad \nu = 1, 2, \ldots, n.$$

One should note that no special assumptions are made about the subsets $\mathscr{C}_\nu$ and $\mathscr{S}_\nu$. They do not need to be polygons, triangles composed of circular arcs, etc. Any two such sets $\mathscr{C}_\nu$ and $\mathscr{S}_\nu$ are said to be congruent if it is possible to transform $\mathscr{C}_\nu$ into $\mathscr{S}_\nu$ by congruence transformations (rotation, reflection, parallel displacement).

At this point it would be natural to consider also the four colour problem as an " unsolved problem ". We have not done this since there have recently been good reports on the state of this problem.

Chapter II

REGULAR PACKINGS IN THE PLANE AND ON THE SPHERE

1. Problems on the Density of Packing

We now formulate in another way problem (**1**) which was stated on p. 1. Let us consider first of all n points P_ν ($\nu = 1, 2, 3, \ldots, n$) distributed arbitrarily on the unit sphere. Let the minimum spherical distance between any two points $P_\nu P_\mu$ be d. Then it is possible to draw about the points P_ν spherical caps of spherical radius $r = \frac{d}{2}$ which have no points in common. By displacing the points P_ν, the minimum distance d is altered and, from continuity considerations, becomes a maximum for (at least) one arrangement of the points P_ν. The surface area of the circles (spherical caps) of radius $r = \frac{d}{2}$ drawn about the points P_ν ($\nu = 1, 2, \ldots, n$) for such an extremal configuration becomes a maximum together with d.

Therefore problem (**1**) is equivalent to problem (**1a**):—

Problem (1a): *n congruent spherical caps must be packed without overlapping on the unit sphere so that they are as large as possible.*

The centres of such an extremal system of caps are then the required capital cities for the dictators of problem (**1**).

Similarly we can reduce problem (**2**) to a covering problem. Let us imagine spherical caps of spherical radius ρ, covering the whole sphere † and placed about the points P_ν which were introduced as fuel depots. Then every point on the sphere is within a spherical distance of at most ρ from one of the points P_ν and our problem is

† i.e. more precisely:—Every point of the sphere must be an interior point *or a boundary point* of at least one of the caps. Therefore in the covering problem we consider the caps to be closed whereas in the packing problem (**1a**) we consider them to be open: caps are allowed to touch each other.

to arrange the covering with the spherical caps so that one obtains a covering of the whole sphcre with caps of the smallest possible spherical radius ρ. Again let us choose the spherical radius to be unity so that (**2**) is equivalent to:

Problem (2a): *The unit sphere must be completely covered by n spherical caps of smallest possible radius.*

The centres of such an extremal system of caps are then the positions for the fuel depots mentioned in problem (**2**).

Now the area of a spherical cap of height h, or of spherical radius r, on the unit sphere is

$$J(r) = 2\pi \,.\, 1 \,.\, h = 2\pi(1 - \cos r).$$

Therefore $J(r)$, and similarly the density of packing

$$D_n = \frac{n \,.\, J(r)}{4\pi} = \frac{n}{2}(1 - \cos r) \tag{1}$$

is a monotonically increasing function of r. Thus the solution of problem (**1a**) yields a system of caps which is packed on the sphere and has *maximal* density of packing, while the solution of (**2a**) gives a covering system of caps of *minimal* density of packing. The density of packing D_n, defined by (1), for a system of caps has an intuitively obvious geometrical meaning: 100 D_n gives the percentage of the surface of the sphere which is covered by the caps. For the packing problem we have $D_n < 1$; for the covering problem, on the other hand, $D_n > 1$.

Before we turn to the solution of problems (**1a**) and (**2a**) for definite natural numbers n, we want to formulate two simpler problems, the treatment of which will be useful for the solution of the original problems:—

Problem (1b): *Let there be n congruent, non-overlapping spherical caps packed on the unit sphere. We must estimate an upper limit $l(n)$, depending on n, for the density of packing* (1) *of this system.*

Problem (2b): *Let the unit sphere be completely covered by n congruent spherical caps. A lower limit $L(n)$ must be estimated for the density of packing of the system.*

We define these limit functions $l(n)$ and $L(n)$ so that, at least for certain natural numbers n, they are *equal* to the density of packing

of a well-defined system of caps. Then the original problem is solved at least *for these numbers n* when the limit functions have been determined.

However, we postpone for the time being the solution of problems (**1b**) and (**2b**) and first of all, as a kind of preparatory exercise, solve the corresponding packing problems *in plane geometry*. This digression is in fact more than a preliminary exercise. We shall see that packing problems in the plane, which are interesting in themselves, lead to problems which are still unsolved (p. 26).

Let us consider therefore how one can pack a system of non-overlapping, unit circles as densely as possible in the plane and, conversely, how one can completely cover a finite area of the plane with as few unit circles as possible. We shall have to give a new definition using a limit for the idea of density of packing in the plane. One cannot avoid introducing a suitable limiting process in plane geometry since the whole plane cannot be covered by a finite number of circles.

Let $\{c_i\}$ be a denumerable set of unit circles in the plane $\mathscr{P}$. Let $\mathscr{C}(R)$ be a circle of radius R about an arbitrary fixed point O of our plane and $n(R)$ the number of circles of $\{c_i\}$ which have an interior point in common with $\mathscr{C}(R)$.

Then

$$\sum_R J_i = \pi \sum_R 1 = \pi \,.\, n(R)$$

is the sum of the areas of these circles.† The limit

$$\lim_{R\to\infty} \frac{\sum_R J_i}{\pi R^2} = \lim_{R\to\infty} \frac{n(R)}{R^2} = \lim_{R\to\infty} \delta_R \tag{2}$$

is called the density of packing of the system $\{c_i\}$.

First of all we note that this definition of the density of packing is independent of the particular choice of point O for the centre of the circle $\mathscr{C}(R)$. Let O^+ be a second point (at a distance d from O) which is the centre of the circle $\mathscr{C}^+(R)$ with radius R. If Σ^+ denotes summation over all the circles of $\{c_i\}$ which have an interior point in common with $\mathscr{C}^+(R)$, then

$$\sum_{R-d} J_i \leqq \sum_R^+ J_i \leqq \sum_{R+d} J_i$$

† Here and in the following, summation over R means that the summation must be taken over all the circles of our system which have an interior point in common with $\mathscr{C}(R)$.

is true and hence

$$\left(1-\frac{d}{R}\right)^2 \frac{\sum\limits_{R-d} J_i}{\pi(R-d)^2} \leqq \frac{\sum\limits_{R}^{+} J_i}{\pi R^2} \leqq \left(1+\frac{d}{R}\right)^2 \frac{\sum\limits_{R+d} J_i}{\pi(R+d)^2}$$

follows; therefore

$$\lim_{R\to\infty} \frac{\sum\limits_{R}^{+} J_i}{\pi R^2} = \lim_{R\to\infty} \frac{\sum\limits_{R} J_i}{\pi R^2}.$$

Now we are able to formulate the problems of plane geometry which correspond to problems (**1b**) and (**2b**):—

Problem (1c): *Let $\{c_i\}$ be a set of unit circles which have no interior point in common in a plane $\mathscr{P}$. An upper bound is required for the density of packing of this system.*

Problem (2c): *In a plane $\mathscr{P}$ let $\{c_i^+\}$ be a set of unit circles with the property that every circle $\mathscr{C}(R)$ about a point $O \in \mathscr{P}$ will be completely covered by finitely many circles of the system $\{c_i^+\}$. It is required to find the lower bound for the density of packing of $\{c_i^+\}$.*

It is very simple to solve these problems if one assumes a *lattice-like* arrangement of the centres [II 3]. However, we want to avoid such particular assumptions and we start the solution by deriving some lemmas which are also important for the solution of further problems.

2. Lemmas

By a plane *polygonal net* we mean a set of simple, closed polygons which lie in a plane and which satisfy the following properties:—

1. No two polygons have an interior point in common.
2. Any two polygons may have vertices or edges in common. If they have edges in common, then they always have one or more *complete* edges in common.
3. All free edges † form a single, simple, closed polygon.

This last condition excludes the possibility of using the term " net " for a system of polygons which lie in a plane and whose union forms a multiply connected domain. On the other hand if, in a case like

† " Free " edges of a net are those which belong to only *one* polygon of the net.

this, one adds to the system the interior boundary polygons, then one has a net in our sense. If we denote the number of polygons of a net by A, the number of their edges (sides) by E and the number of vertices by V, the number

$$C(\mathcal{N}) = V-E+A$$

is called Euler's characteristic number for the net $\mathcal{N}$. We now prove

Theorem 1: *Euler's characteristic number for a plane net is equal to* 1:—

$$C(\mathcal{N}) = V-E+A = 1. \tag{3}$$

For the proof we note first that the characteristic number does not change if one draws an (interior) diagonal in a polygon. When

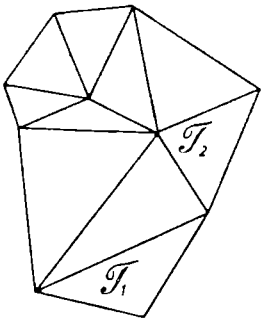


FIG. 1

this is done the number of vertices remains unchanged while A and E each increase by 1. By drawing diagonals like this one can eventually triangulate all the polygons of the net ([II 1] p. 66) and it therefore suffices to prove the theorem for a triangular net.

Thus let $\mathcal{N}_n$ be a net made up of n triangles (fig. 1). We call the triangles of $\mathcal{N}_n$ which have free edges boundary triangles of the net (e.g. $\mathcal{T}_1$ and $\mathcal{T}_2$ in fig. 1). If among the boundary triangles of the net one occurs with two free edges (e.g. $\mathcal{T}_1$ in fig. 1) then, by separating off this triangle, one obtains a system of $n-1$ triangles that similarly fulfil all the conditions required in the definition of a net. It has the same characteristic number as $\mathcal{N}_n$ i.e. $C(\mathcal{N}_{n-1}) = C(\mathcal{N}_n)$, for V and A have decreased by 1 and E by 2. If in the net only triangles with *one* free edge occur,† one can also exclude a triangle in this case without disturbing the connectivity of the

† That is the case e.g. in net $\mathcal{N}_{n-1}$ in Fig. 1 which is formed from $\mathcal{N}_n$ by removing $\mathcal{T}_1$.

remaining system of triangles. If this is done, E and A change by 1 and V not at all.

One can reduce the net in this way until it consists of only one triangle. And, since the characteristic number of a triangle is equal to $3-3+1 = 1$, our theorem is proved. Obviously it is also valid if the sides of the polygons are replaced by continuous (non-intersecting) arcs.

Similarly one can construct on the sphere (say from great circle arcs) polygonal nets. However, in this case there are also systems of polygons which cover the whole sphere. We denote these systems, which are important for the geometry of the sphere, as nets on the sphere. One obtains the exact definition by varying the definition for plane nets. Plane polygons are replaced of course by polygons on the sphere and 3. is to be replaced by the rule:— There are no free edges.

Now let us consider a net like this on the sphere mapped onto the plane by stereographic projection (see e.g. [II 4]). Let the centre for the projection be an interior point of a spherical polygon. Then one obtains a plane net as the plane image of the spherical net. However, care must be taken *not* to count the exterior of the polygon with free edges as a plane polygon, though it is the image of a spherical polygon, namely the image of the polygon which contains the North pole of the stereographic projection. Thus there exist between the vertices, edges and areas (v, e and a) of the spherical net and the corresponding numbers (V, E and A) of the plane net the following relations:—

$$v = V,\ e = E,\ a = A+1.$$

Hence from (3) follows

Theorem 2: *Euler's characteristic number for a spherical net is equal to* 2:

$$v-e+a = 2. \tag{4}$$

A large number of geometrical extremal problems can be solved by use of Jensen's inequality for convex functions. A function $f(x)$ which is defined on an interval $[a, b]$ of the x-axis is called *convex* if, for any two points x_1 and x_2 of the interval,

$$2f\left(\frac{x_1+x_2}{2}\right) \leqq f(x_1)+f(x_2) \tag{5}$$

is always true. It is called *concave*, if, conversely, for any two points x_1, x_2,

$$2f\left(\frac{x_1+x_2}{2}\right) \geqq f(x_1)+f(x_2). \tag{5'}$$

Now for these functions we have

Theorem 3 (Jensen's Inequality): *If $f(x)$ is convex in $[a, b]$ then, for any points $x_\nu \in [a, b]$ $(\nu = 1, 2, 3, \ldots, n)$,*

$$f(x_1)+f(x_2)+\ldots+f(x_n) \geqq n f\left(\frac{x_1+x_2+\ldots+x_n}{n}\right). \tag{6}$$

On the other hand for concave functions $f(x)$

$$f(x_1)+f(x_2)+\ldots+f(x_n) \leqq n f\left(\frac{x_1+x_2+\ldots+x_n}{n}\right). \tag{6'}$$

It is sufficient to prove (6). The proof for concave functions is similar. We show first of all that the theorem is true for $n = 2^m$ (m an integer greater than 1) then that, if the theorem is true for n, it follows that it is also true for $n-1$. Hence the validity of (6) for all natural numbers N is assured. In fact, if a number N lies between 2^{m-1} and 2^m, then, since (6) is valid for 2^m, it follows that it is valid for 2^m-1, hence that it is valid for 2^m-2 and finally that it is true for N.

For $n = 4$ we have first of all by repeated application of (5)

$$\tfrac{1}{4}(f(x_1)+f(x_2)+f(x_3)+f(x_4)) \geqq \tfrac{1}{4}\left(2f\left(\frac{x_1+x_2}{2}\right)+2f\left(\frac{x_3+x_4}{2}\right)\right)$$

$$\geqq f\left(\frac{\frac{x_1+x_2}{2}+\frac{x_3+x_4}{2}}{2}\right) = f\left(\frac{x_1+x_2+x_3+x_4}{4}\right).$$

The validity of (6) for $n = 2^m$, $m>2$, is deduced by a standard induction argument. It only remains to show that, if it is true for n, it follows that it is true for $n-1$. Therefore, let $x_1, x_2, \ldots, x_{n-1}$ be any points of our integral $[a, b]$. The validity of (6) will be proved for these points assuming that the inequality for any n points of the interval is true. We add to the $n-1$ points another point

$$x_n = \frac{x_1+x_2+\ldots+x_{n-1}}{n-1}, \tag{7}$$

which also lies in the interval $[a, b]$. By our induction hypothesis (6) is true for n points. We write this inequality in the form

$$f(x_1)+f(x_2)+\ldots+f(x_{n-1})\geqq n f\left(\frac{x_1+x_2+\ldots+x_n}{n}\right)-f(x_n). \quad (8)$$

If one now multiplies (7) by $n-1$ it becomes clear that the following representation is also true for x_n:

$$x_n=\frac{x_1+x_2+\ldots+x_n}{n}. \quad (7')$$

From (8), using (7') and (7),

$$f(x_1)+f(x_2)+\ldots+f(x_{n-1})\geqq n f(x_n)-f(x_n)=$$
$$(n-1)f\left(\frac{x_1+x_2+\ldots+x_{n-1}}{n-1}\right).$$

This is the inequality we wanted to prove.

We apply Jensen's inequality to prove the following theorem which will be used later:—

Theorem 4: *Of all the n-gons inscribed in a circle of radius r the one with the greatest area is the regular n-gon.*

Let us denote (fig. 2) by $2\alpha_\nu$ the angles subtended at the centre C by the sides A_1A_2, A_2A_3, ... of the polygon $A_1A_2A_3\ldots A_n$ when it is subdivided into isosceles triangles by the radii CA_1, CA_2, ... Then the area of the whole polygon is

$$A=r^2\sum_{\nu=1}^{n}\sin\alpha_\nu\cos\alpha_\nu=\frac{r^2}{2}\sum_{\nu=1}^{n}\sin 2\alpha_\nu.$$

Now, in the interval $[O, \pi]$, $\sin x$ is a concave function. By a well-known addition formula

$$\sin x_1+\sin x_2=2\sin\frac{x_1+x_2}{2}\cos\frac{x_1-x_2}{2}\leqq 2\sin\frac{x_1+x_2}{2}.$$

Thus, by Jensen's theorem (6'), it is true that

$$A\leqq\frac{r^2}{2}n\sin\frac{2\alpha_1+\ldots+2\alpha_n}{n}=\frac{nr^2}{2}\sin\frac{2\pi}{n}.$$

On the right hand side of this inequality we have just the area of the regular n-gon inscribed in the circle.

In this case we used the addition formula to prove the concave

character of $\sin x$. In many cases it is convenient to use the following criterion:—

A function $f(x)$ which is twice continuously differentiable in the interval $[a, b]$, is convex (concave) if the second derivative $f''(x)$ is positive (negative) in the whole interval.

Since this important criterion is not proved in many text-books on Analysis we include the proof here. By Taylor's theorem

$$f(\xi) = f(x)+f'(x)(\xi-x)+\tfrac{1}{2}f''(x+\vartheta(\xi-x))(\xi-x)^2 \qquad (9)$$

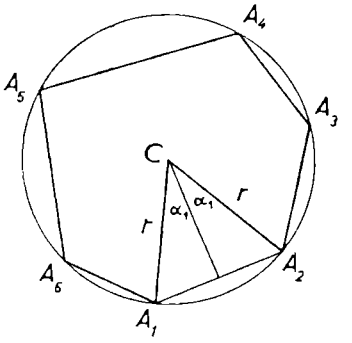


FIG. 2

is true for $x \in [a, b]$, $\xi \in [a, b]$. Now let $f''(x)>0$ in the whole interval. Then from (9) it follows that

$$f(\xi)>f(x)+f'(x)(\xi-x).$$

Hence, for $\xi = x_1$ and $\xi = x_2$, respectively,

$$f(x_\nu)>f(x)+f'(x)(x_\nu-x) \quad (\nu = 1, 2),$$

and therefore, if $x_1+x_2 = 2x$,

$$f(x_1)+f(x_2)>2f(x)+f'(x)(x_1+x_2-2x) = 2f(x).$$

Similarly, one gets (5′) if $f''(x)$ is negative in $[a, b]$.

3. PACKING PROBLEMS IN THE PLANE

Let $\{c_i\}$ be a set of non-overlapping unit circles in a plane. To estimate the density of packing of such a system (problem **(1c)**) we can suppose without any loss of generality that the system is " saturated ", i.e. it is not possible to pack another unit circle in the plane without some overlapping.

We now join the centres of the circles by straight lines and so form a triangular net whose vertices are all centres of the circles. This net can be constructed so that there is no angle greater than

120° in any of the triangles. Let us begin to construct the net by joining any one of the centres C_1 with the closest (or one of the closest) centres C_2. Now we consider the circles having interior points in common with the parallel strip formed by the lines through C_1 and C_2 perpendicular to C_1C_2 (fig. 3). There must exist circles with this property, otherwise the system would not be saturated.

The number 3 is given to the circle of this set whose centre is closest to the straight lines. If there happen to be several of these circles one is chosen arbitrarily. If its centre lies within the parallel

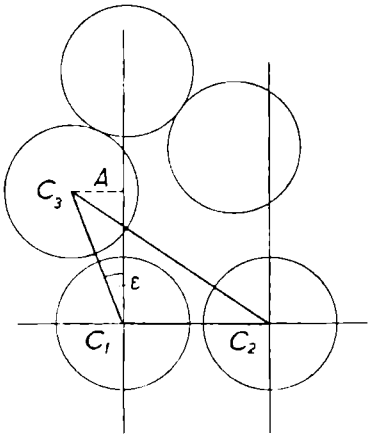


FIG. 3

strip, then the triangle $C_1C_2C_3$ is acute angled. However, if C_3 lies outside the strip and if c_3 is cut by the perpendicular through C_1, say, (fig. 3) then, for the angle ε between C_1C_3 and the perpendicular, it is true that

$$\sin \varepsilon = \frac{C_3A}{C_1C_3} < \tfrac{1}{2}.$$

The angle $C_3C_1C_2$ is therefore less than 120°. The net can now be constructed by adjoining further triangles so that there is no angle greater than 120° in any of the triangles. Then each triangle in the net has an angle α for which $60° \leqq \alpha < 120°$ is true. Since the sides of these triangles are at least of length 2, we get for the area $J(\mathcal{T})$ of each of these triangles

$$J(\mathcal{T}) \geqq \tfrac{1}{2} \cdot 2 \cdot 2 \cdot \sin \alpha \geqq \sqrt{3}. \tag{10}$$

We have equality in (10) only for triangles $C_\lambda C_\mu C_\nu$ whose circles c_λ, c_μ, c_ν touch each other in pairs. For later use we also note a (quite crude) upper bound of the area of the triangle. Since the system is saturated it follows that the base and height of the triangles

must be less than 4. Therefore, we have

$$J(\mathscr{T})<8. \tag{10'}$$

Now let us draw the triangular net of the centres of all the circles which have interior points in common with a circle $\mathscr{C}(R)$ of radius R and centre O. Then for the quotient δ_R which arose in definition (2) for the density of packing it is true that

$$\delta_R = \frac{V\pi}{\pi R^2} \leqq \frac{V\pi}{A\,.\,\sqrt{3}} \cdot \frac{J(\mathscr{N})}{\pi R^2}, \tag{11}$$

where V is the number of vertices of our net, A the number of areas and $J(\mathscr{N})$ the area of the whole net for which we have from (10) and (10')

$$A\,.\,\sqrt{3} \leqq J(\mathscr{N}) < 8A. \tag{10''}$$

We also note that the second fraction on the right side of (11) converges to 1 as R increases. From the definition of the net it follows that the distances of its vertices from the centre O of $\mathscr{C}(R)$ are less than $R+1$. Since the system $\{c_i\}$ is saturated

$$\pi(R-3)^2 < J(\mathscr{N}) < \pi(R+1)^2, \tag{10'''}$$

and hence

$$\lim_{R\to\infty} \frac{J(\mathscr{N})}{\pi R^2} = 1. \tag{12}$$

Next we express the number of areas A in (11) in terms of V using Theorem 1. To do this the edges of our net in each of the triangles are enumerated including all the free edges once and all the others twice. Therefore we have

$$3A = 2E - E',$$

where E' indicates the number of free edges. Thus it follows from (3) that

$$A = 2V - E' - 2 \tag{13}$$

and hence (11) becomes

$$\delta_R \leqq \frac{\pi V}{(2V-E'-2)\sqrt{3}} \cdot \frac{J(\mathscr{N})}{\pi R^2}. \tag{14}$$

We will immediately get the desired bound for the density of packing from (14) when we have proved that

$$\lim_{R\to\infty} \frac{E'}{V} = 0. \tag{15}$$

For this proof we note that E' is not only the number of edges but also the number of angles of the boundary polygon. The area of the circles corresponding to these vertices must, however, be smaller than the area of the ring with radii $R+2$ and $R-4$ in which the circles lie. Therefore

$$E'\pi < \pi(R+2)^2 - \pi(R-4)^2 = 12\pi(R-1). \tag{16}$$

Further, from (13), we have $2V > A$, and thus it follows from (16), taking (10″) and (10‴) into account that

$$\frac{E'}{V} < \frac{12\pi(R-1)\,.\,2}{A} < \frac{192\pi(R-1)}{J(\mathcal{N})} < \frac{192(R-1)}{(R-3)^2}.$$

(15) immediately follows and from (14), (15) and (12) we obtain

$$d = \lim_{R\to\infty} \delta_R \leqq \lim_{V\to\infty} \frac{\pi}{\left(2 - \frac{E'}{V} - \frac{2}{V}\right)\sqrt{3}} \,.\, 1 = \frac{\pi}{2\sqrt{3}}.$$

Hence we have proved:—

Theorem 5: *The density of packing of a system* $\{c_i\}$ *of non-intersecting unit circles in the plane is at most equal to* $\frac{\pi}{2\sqrt{3}}$:

$$d = \lim_{R\to\infty} \frac{\pi V}{\pi R^2} \leqq \frac{\pi}{2\sqrt{3}} = 0{\cdot}9069.... \tag{17}$$

It is easy to satisfy oneself that this bound is attained for a system of circles whose net contains only equilateral triangles with sides of length 2 (fig. 4).

To solve problem (**2c**) (p. 8) we again construct a triangular net from the centres of the system of unit circles. To do this we must take all the centres of the circles in $\{c_1^+\}$ which have at least one interior point in common with a circle $\mathfrak{C}(R)$ about O. This is done as follows:—Let C_1 and C_2 be the centres of two such circles c_1^+ and c_2^+ with non-empty intersection. One vertex P of the common lune must then be covered by a further circle c_3^+. Its centre is denoted by C_3. If P is covered by several circles then any one of them is chosen. Now, by the same method, further triangles are added to the triangle $C_1C_2C_3$ so that finally all the centres in question are included.†

† It is possible that after completing the process of adding on triangles there are still some centres of circles which lie *in the interior* of the triangles. By joining lines to the sides of triangles already present these centres are subsequently attached to the net.

In this way we obtain a triangular net in which *every triangle has a circumcircle of radius at most equal to* 1. The area of each of these triangles is therefore, by Theorem 4, at most equal to the area of the equilateral triangle inscribed in the unit circle:—

$$J(\mathscr{T}) \leqq \tfrac{3}{4}\sqrt{3}.$$

Therefore for the area of the whole net it is true that

$$J(\mathscr{N}) \leqq A \,.\, \tfrac{3}{4}\sqrt{3}.$$

However, the number A of areas of the net is, by (13), less than $2V$ and hence we have for the density of packing, taking (12) into account,

$$D = \lim_{R\to\infty} \frac{\pi V}{\pi R^2} = \lim_{R\to\infty} \frac{\pi V}{J(\mathscr{N})} \,.\, \frac{J(\mathscr{N})}{\pi R^2} \geqq \frac{\pi V \,.\, 4}{3\sqrt{3} \,.\, 2 \,.\, V} \,.\, 1.$$

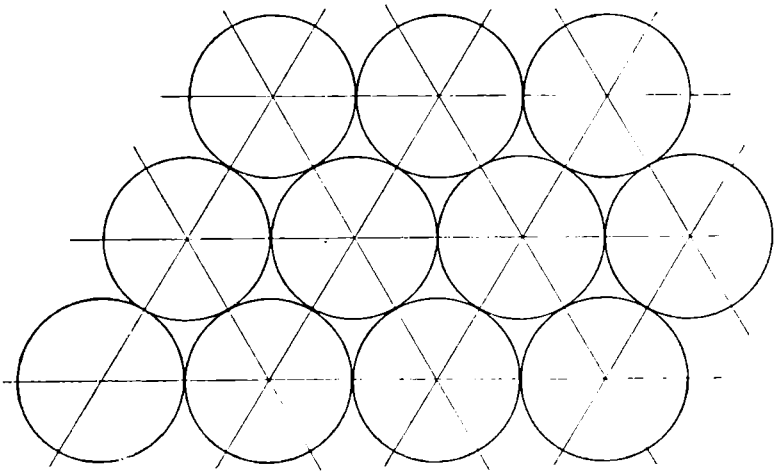


FIG. 4

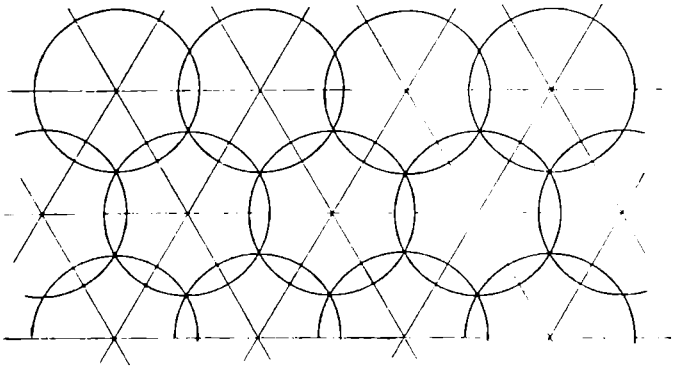


FIG. 5

We sum up in

Theorem 6: *Let $\{c_i^+\}$ be a set of unit circles in a plane $\mathscr{P}$ with the property that every circle $\mathscr{C}(R)$ with centre $O \in \mathscr{P}$ is covered by finitely many circles of the system $\{c_i^+\}$. Then we have for the density of packing of this system*

$$D = \operatorname*{lim}_{R\to\infty} \frac{\pi V}{\pi R^2} \geqq \frac{2\pi}{3\sqrt{3}} = 1{\cdot}209.... \tag{18}$$

In this case also the extremal value is attained when we have a packing for which the net of centres contains only equilateral triangles (fig. 5).

4. Packing Problems on the Sphere

To solve problems (**1b**) and (**2b**) we consider, as for plane problems, a triangular net constructed from the centres of caps. In the covering problem (problem (**2b**)) this net can be formed so that the spherical radius of the circumscribed circle of each triangle is at most equal to the radius r of the caps of the system. To construct this net we follow a procedure that is completely analogous to that for the corresponding plane problem (p. 13).

Then the surface area of a triangle of a net like this is at most equal to that of an equilateral triangle inscribed in a spherical circle of radius r. The proof of this extremal property of the triangles of the net could be carried out using the Jensen inequality. Instead we want to use the theorem about Lexell's circle:—

Theorem 7: *The geometrical locus of the apexes of all spherical triangles with equal surface area which have a side AB in common with a given triangle ABC, is the spherical circle passing through C and the antipodal points A' and B' of A and B.*

To prove this theorem we drop from A, B and C onto the great circle passing through the midpoints D and E of AC and BC (the "middle-line" m) the perpendiculars which meet m in G and H, say, (fig. 6). Then the following triangles are congruent:

$$\Delta CDF \equiv \Delta ADG, \quad \Delta CEF \equiv \Delta BEH,$$

since in each case they have one side and two corresponding angles equal. Therefore the triangle ABC is equal in area to the quadrangle $ABHG$ (right-angled at G and H).† Let C_1 be another point on the

† It is easy to see that this and the following statements are also valid for obtuse-angled triangles ABC.

hemisphere which is determined by m and contains C and with its perpendicular C_1F_1 onto m equal to $CF = AG = HB = d$. Then, simply by considering congruence, it is shown that m is also the middle-line for the triangle ABC_1. Therefore triangle ABC_1 has also the same surface area as the quadrangle $ABHG$ and triangle ABC. The geometric locus for all points which are equidistant from a great circle on the sphere, is, however, a spherical circle a, the " distance line ".†

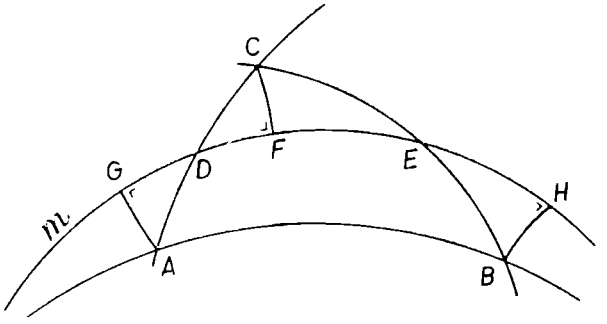


FIG. 6

All triangles ABC^+ whose apexes C^+ lie on a, have therefore equal surface area and it is obvious that, conversely, triangles ABC^{++} with C^{++} *not* lying on a have a smaller or greater area than triangle ABC. It only remains to determine the position of the distance line. Obviously the points A' and B' which are antipodal to A and B are the same distance d from the great circle m as A and B. Since A' and B' lie on the hemisphere determined by m and containing C, the distance line a must also pass through A' and B'. This completes the proof.

Now let ABC be an isosceles triangle inscribed in a circle $\mathscr{A}$ and with apex C. Then the Lexell circle passing through C must, by symmetry, touch $\mathscr{A}$ and clearly *on the outside*. This means, however, that among all the triangles inscribed in circle $\mathscr{A}$ and with base AB the isosceles triangle has the greatest surface area. If ABC is not *equilateral* then one can immediately find among the triangles inscribed in $\mathscr{A}$ a greater one choosing AC as base. From continuity considerations there must be a greatest triangle among all those

† If one calls the given great circle " the equator " of the sphere, then the distance line is a circle of latitude (not a great circle!) lying in the northern or southern hemisphere.

inscribed in $\mathscr{A}$. Hence the equilateral triangle must be this greatest triangle. We have therefore:

Theorem 8: *Of all the triangles inscribed in a spherical circle the equilateral triangle has the greatest area.*

Using this theorem we can now estimate the density of packing for the given system of caps (problem (**2b**)). By Theorem 8 the area of the triangles of the net is at most equal to the surface area S_m of an equilateral triangle inscribed in the circle of spherical radius r. Therefore let us calculate S_m.

Let LMN be an equilateral triangle with circumscribed circle of radius r and angle 2α. Let U be the centre of the circumscribed circle and O the foot of the perpendicular from U onto LM. Then the triangle LUO has at U the angle $\pi/3$, at O the angle $\pi/2$ and

$$\cot \alpha = \cos r \,.\, \tan \frac{\pi}{3},$$

therefore

$$S_m = 6\alpha - \pi = 6 \cot^{-1} (\sqrt{3} \cos r) - \pi. \tag{19}$$

Now let v, e and a again be the numbers of vertices, edges and areas of our spherical net. Since it covers the whole sphere, we have the surface area of this net

$$4\pi = S(\mathscr{N}) \leqq a \,.\, S_m. \tag{20}$$

Since the net consists only of triangles then $3a = 2e$ and from (4) it follows that

$$a = 2v - 4. \tag{21}$$

Substituting (19) and (21) into (20) it follows that

$$6 \cot^{-1} (\sqrt{3} \cos r) - \pi \geqq \frac{4\pi}{2v-4},$$

or

$$\cos r \leqq \frac{1}{\sqrt{3}} \cot \frac{v\pi}{3(2v-4)}. \tag{22}$$

Now we again replace the number of vertices v by the (equal) number n of spherical caps of our system and from (22) obtain for the density of packing defined by (1):—

Theorem 9: *For the density of packing D of a system of n congruent spherical caps covering the unit sphere*

$$D \geqq \frac{n}{2}\left(1 - \frac{1}{\sqrt{3}} \cot \omega_n\right), \quad \omega_n = \frac{n\pi}{6(n-2)}. \tag{23}$$

There is equality in (22) and (23) only when all the triangles of the net are equilateral. However, this is possible only if the n centres are the n vertices of a regular triangular polyhedron inscribed in the sphere. Therefore, there exist such (regular) packings (by the known theorems about regular polyhedra, see e.g. [II 7]) for the numbers $n = 3, 4, 6$ and 12. The bound (23) can be attained by a suitable packing only in these cases. We set down in a table the values for the smallest radius of the cap and the corresponding density of packing which arise in these cases:—

n	3	4	6	12
$\cos r$	0	$\frac{1}{3}$	$\frac{1}{3}\sqrt{3}$	$\frac{1}{3}\sqrt{3+\frac{6}{5}\sqrt{5}}$
r	90°	70° 52′	54° 46′	37° 43′
D	1·5	1·33...	1·268...	1·254...

Hence the problem (**2**) is also solved for the cases $n = 3, 4, 6$ and 12. The situations for the fuel depots are the vertices of a regular triangular polyhedron with the corresponding number of vertices.

Theorem 9 was first proved by FEJES TOTH in 1943 [II 2] (by other methods).

To solve the problem (**1a**) (p. 5) one could try to carry over the procedure used in the proof of Theorem 5. However, we choose to describe as many different kinds of methods of solution as possible and therefore the theorem of Jensen will be used this time. Thus, let there be given a system of n congruent, non-overlapping spherical caps of spherical radius r on the unit sphere. We assume that the system is saturated and that therefore no more caps of radius r can be packed without overlapping. Then the tangential planes touching the sphere at the centres C_ν ($\nu = 1, 2, 3, \ldots, n$) of the caps form a polyhedron whose central projection on the sphere produces a net $\mathcal{N}$ on it. If one joins the centres of the caps by arcs of great circles which intersect the edges of the net $\mathcal{N}$ at right-angles then one obtains a second net $\mathcal{N}'$ which contains only triangles (fig. 7 shows the corresponding picture in the plane).

Let v, e and a be as usual the numbers of vertices, edges and areas of the first net $\mathcal{N}$, v', e' and a' the corresponding numbers for $\mathcal{N}'$. Then it is obvious that $e = e'$, $v = a'$, $a = v'$. Since $\mathcal{N}'$ is a triangular net, we have moreover that

$$2e' = 2e = 3a' = 3v. \tag{24}$$

The polygons of $\mathcal{N}$ can now be decomposed into isosceles spherical triangles by joining the n centres to the vertices of the

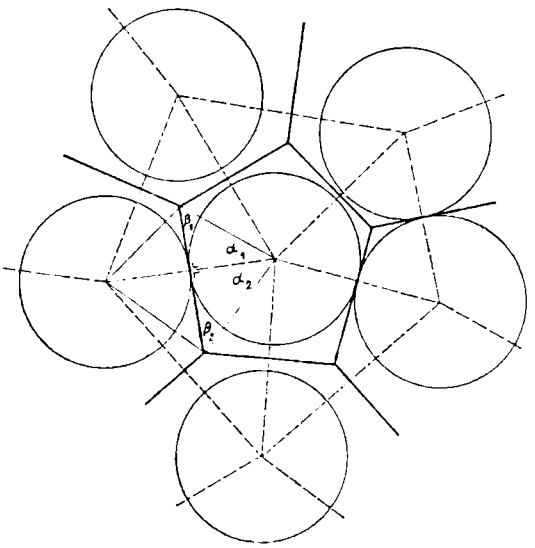


FIG. 7

polygons. These triangles are then divided by the sides of $\mathcal{N}'$ into right-angled triangles so that at each of the e edges of $\mathcal{N}$ there are 4 right-angled triangles (fig. 7). In the right-angled triangles let α_ν be the angles at the centres, β_ν the angles at the vertices of the polygons and b_ν the sides opposite the β_ν. Then the area of the right-angled triangle with index ν is:—

$$\Delta(\alpha_\nu) = \alpha_\nu + \beta_\nu - \frac{\pi}{2}.$$

Since

$$\sin\left(\frac{\pi}{2} - \beta_\nu\right) = \cos b_\nu \,.\, \sin \alpha_\nu$$

it follows that

$$\Delta(\alpha_\nu) = \alpha_\nu - \sin^{-1}(\cos b_\nu \sin \alpha_\nu). \tag{25}$$

Since $b_\nu \geqq r$, we have for all ν

$$\Delta(\alpha_\nu) \geqq \alpha_\nu - \sin^{-1}(\cos r \sin \alpha_\nu).$$

Now the $4e$ triangles just cover the whole sphere. Thus

$$4\pi \geqq \sum_{\nu=1}^{4e} (\alpha_\nu - \sin^{-1}(\cos r \sin \alpha_\nu)) = \sum_{\nu=1}^{4e} s(\alpha_\nu) \tag{26}$$

is true, where the function

$$s(\alpha) = \alpha - \sin^{-1}(\cos r \sin \alpha)$$

gives, by (25), the area of a right-angled spherical triangle with one angle α and the adjacent side r. Now, using the criterion on p. 13, one immediately establishes that $s(\alpha)$ is a convex function for $0 \leqq \alpha \leqq \pi/2$. Of course, $s''(\alpha)$ is positive as one can immediately verify.† Applying Jensen's theorem (Theorem 3), it follows from (26) that

$$4\pi \geqq 4e\, s\left(\frac{\pi n}{2e}\right) = 4e\left[\frac{\pi n}{2e} - \sin^{-1}\left(\cos r \sin \frac{\pi n}{2e}\right)\right], \tag{27}$$

since

$$\alpha_1 + \alpha_2 + \ldots + \alpha_{4e} = 2\pi \,.\, n.$$

However, from (24) and (4)

$$e = 3(a-2) = 3(n-2)$$

and hence from (27) we have

$$4\pi \geqq 12(n-2)\left[\frac{\pi n}{6(n-2)} - \sin^{-1}\left(\cos r \,.\, \sin \frac{\pi n}{6(n-2)}\right)\right]$$

or

$$\frac{\pi}{6} \leqq \sin^{-1}\left(\cos r \,.\, \sin \frac{\pi n}{6(n-2)}\right),$$

therefore

$$\cos r \geqq \frac{1}{2 \sin \omega_n}, \quad \omega_n = \frac{\pi n}{6(n-2)}. \tag{28}$$

If one substitutes this estimate into the definition of the density of packing, one has at once

Theorem 10: *For the density of packing*

$$d_n = \frac{n}{2}(1 - \cos r)$$

† One can immediately derive the convexity of $s(\alpha)$ from the geometrical meaning of this function. The details of the reasoning are left to the reader.

of a system of n congruent, non-overlapping spherical caps of spherical radius r

$$d_n \leqq \frac{n}{2}\left(1-\frac{1}{2\sin\omega_n}\right), \quad \omega_n = \frac{\pi n}{6(n-2)} \tag{29}$$

is true.

Here too the equality sign is possible only when all the triangles of the net $\mathcal{N}'$ are equilateral. If Ω_n is the half-angle and $2r$ the side of such a spherical triangle then we have

$$\sin\Omega_n = \frac{\sin r}{\sin 2r} = \frac{1}{2\cos r}, \tag{30}$$

therefore

$$\cos r = \frac{1}{2\sin\Omega_n}.$$

This angle Ω_n is just equal to the angle ω_n of (28). For the area of our equilateral triangle is $J(\Delta) = 6\Omega_n - \pi$. If $a' = 2(n-2)$ of such triangles cover the whole sphere, we have

$$2(n-2)(6\Omega_n-\pi) = 4\pi, \quad \Omega_n = \omega_n.$$

The converse is left for the reader to prove. Thus we have given a geometrical interpretation for the angle ω_n appearing in (23) and (29) when n is the number of vertices of a regular triangular polyhedron.

We note further that, when the *inequality* holds in (28) and (29),

$$\omega_n = \frac{\pi n}{6(n-2)} > \Omega_n. \tag{31}$$

In fact, in this case we have

$$\sin\omega_n > \frac{1}{2\cos r} = \sin\Omega_n,$$

therefore $\omega_n > \Omega_n$.

Again we collect the results in a table for the cases $n = 3, 4, 6$ and 12:—

n	3	4	6	12
$\cos r$	$\frac{1}{2}$	$\frac{1}{3}\sqrt{3}$	$\frac{1}{2}\sqrt{2}$	0·852
r	60°	54° 46′	45°	31° 43′
d	0·75	0·845	0·878	0·896

Thus the problem (1) is also solved for some values of n. In the cases $n = 3, 4, 6$ and 12, just as for problem (2), the vertices of a regular triangular polyhedron inscribed in the sphere form the point system which is the solution of the problem. We shall see in the next chapter that for other values of n the solutions of the two problems are by no means always the same.

Another form of Theorem 10 is important for some problems:—

Theorem 10a: *Given n $(n>3)$ points on the unit sphere there is always a pair of points whose distance apart is*

$$\delta_n \leqq \sqrt{4-\frac{1}{\sin^2 \omega_n}}. \tag{32}$$

By the distance δ_n between two points P and Q on the sphere we mean here the length of the line-segment PQ. The spherical distance corresponding to δ_n is $\delta_n^* = 2 \sin^{-1} \frac{\delta_n}{2}$.

The bound (32) can easily be derived from (28). Obviously from (28) it follows that

$$\sin^2 \omega_n \geqq \frac{1}{4 \cos^2 r},$$

therefore

$$\delta_n^* = 2r \leqq \cos^{-1}\left(\frac{\cot^2 \omega_n - 1}{2}\right).$$

From this we calculate the limit given in (32). In fact

$$1-2 \sin^2 \frac{\delta_n^*}{2} = \cos \delta_n^* \geqq \frac{\cot^2 \omega_n - 1}{2},$$

therefore

$$-4 \sin^2 \frac{\delta_n^*}{2} \geqq \frac{1}{\sin^2 \omega_n} - 4$$

or

$$\delta_n = 2 \sin \frac{\delta_n^*}{2} \leqq \sqrt{4-\frac{1}{\sin^2 \omega_n}}.$$

5. Further Problems

To conclude this chapter we examine some unsolved problems which are closely related to those already considered.

A packing of non-overlapping circles in the plane † is said to be *rigid* if every circle is touched by at least three other circles so that the points of contact on any of the circles are less than a semi-circle apart. The packing shown in fig. 4 is certainly rigid: for every circle is enclosed by six others and every circle need only be touched by three others as described above to make a displacement of the circles (without overlapping) impossible.

A natural example of a packing of circles in the plane in which every circle is touched by three others is shown in fig. 8. The lines from one centre to the centres of the neighbouring circles form the edges of a trihedral whose arms enclose angles of 120°. All the edges of the trihedrals form a net of hexagons. From this observation we immediately obtain the density of packing of the system. In each hexagon (e.g. $P_1P_2P_3P_4P_5P_6$ in fig. 8) there are six circular sectors with angle 120° and therefore two complete circles. Hence the density of packing ‡ of the system is

$$d_1 = \frac{2\pi \cdot 1}{6 \cdot \frac{4}{4}\sqrt{3}} = \frac{\pi}{3\sqrt{3}} \sim 0{\cdot}6046. \tag{33}$$

However, this system $\mathscr{S}_1$ has by no means the smallest density of packing of all rigid systems. One immediately obtains a system $\mathscr{S}_2$ with smaller density if one replaces every circle of $\mathscr{S}_1$ by three congruent circles which touch the first circle at its points of contact (see $\mathscr{C}_1$ in fig. 8). If, in addition, one requires that every two of the three smaller circles must touch each other, then the new system $\mathscr{S}_2$ of smaller circles is uniquely determined. It is also rigid in the sense of our definition since every circle $\mathscr{C}_2 \in \mathscr{S}_2$ is touched by three others, namely by two which belong to the same circle of the system $\mathscr{S}_1$ as $\mathscr{C}_2$ and by one from the neighbouring circle. In fig. 8 the net of the centres of $\mathscr{S}_2$, consisting of triangles and dodecagons is drawn in part of the figure. The edges of the trihedrals of this system enclose an angle of 60° and two angles of 150°.

† Of course the corresponding definition is possible for a system of non-overlapping spherical caps on the sphere.

‡ It follows from an estimate, obtained similarly to that on p. 15, that the density of packing as defined in (2) can be evaluated in this way from the covering of the single hexagonal cell.

We calculate the radius ρ of the circles $\mathscr{S}_2$, knowing the radius $r = 1$ for the large circles, as follows:—

$$\tfrac{1}{2}\sqrt{3} = \sin 60° = \frac{\rho}{r-\rho},\ \rho = r(2\sqrt{3}-3) \sim 0{\cdot}464\ .\ r. \qquad (34)$$

Therefore, since three circles of $\mathscr{S}_2$ only cover the fraction $\dfrac{3\pi\rho^2}{\pi}$ of a circle $\mathscr{C}_1 \in \mathscr{S}_1$, we get for the density of packing of $\mathscr{S}_2$ from (33) and (34)

$$d_2 = 3\rho^2 d_1 = (7\sqrt{3}-12)\pi \sim 0{\cdot}393.$$

Probably this is the smallest possible value for the density of a rigid packing of circles. In any case it is not possible to repeat the same game of replacing a circle by three smaller ones. We know that the angles formed by the edges of the trihedrals of $\mathscr{S}_2$ are 60°, 150°, 150°. Therefore it is not possible to place at each of the three points

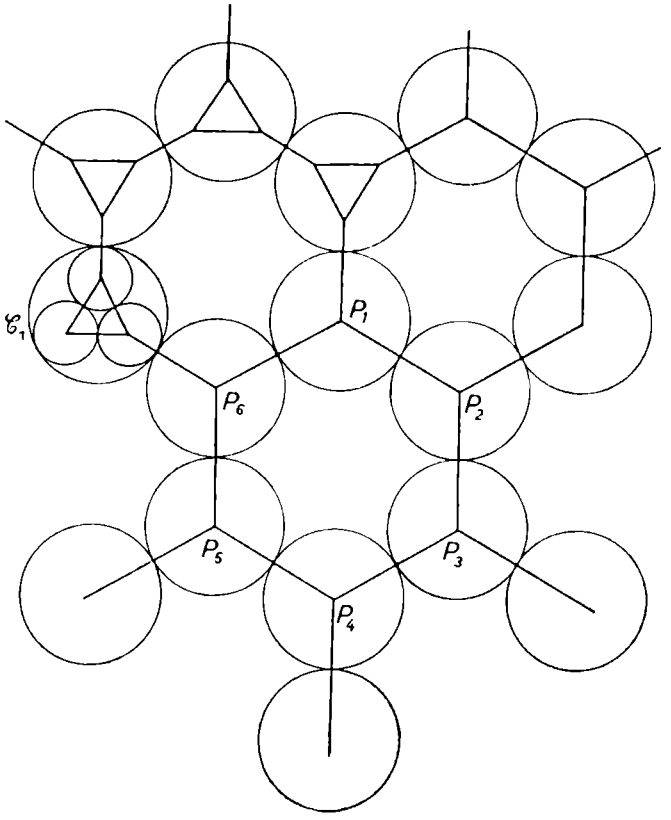


FIG. 8

of contact of an $\mathscr{S}_2$ circle one of three congruent circles touching the $\mathscr{S}_2$ circle and touching each other in pairs.

Other unsolved problems are referred to briefly in a publication of A. HEPPES [II 11]. If every point of the plane is covered by *at most* k (congruent) circles, then we speak of a k-fold packing of circles. If every point belongs to *at least* k circles then there is a k-fold covering of the plane. Let the corresponding densities of packing be δ_k and Δ_k and the extremal values

$$d_k = \overline{\lim}\, \delta_k, \quad D_k = \underline{\lim}\, \Delta_k.$$

Then from our results we have

$$d_1 = \frac{\pi}{\sqrt{12}}, \quad D_1 = \frac{2\pi}{\sqrt{27}}.$$

Further it is obviously true that

$$kd_1 \leqq d_k < k, \quad kD_1 \geqq D_k > k. \tag{35}$$

One obtains, for example, a k-fold packing of circles of density kd_1 when one replaces each circle in the simple extremal packing by k copies placed over one another. However, it is easy to see that there exist k-fold packings of circles for which the density will be greater that kd_1. A. HEPPES gives a simple example of this in his work. Hence the equality sign cannot hold in the first inequality in (35).

There still remains the open question:—

How large are the numbers

$$d_k = \overline{\lim}\, \delta_k, \quad D_k = \underline{\lim}\, \Delta_k$$

for $k > 1$*?*

Chapter III

IRREGULAR PACKINGS AND COVERINGS

1. The Existence Theorem

Let us now deal with the solution of problems (**1a**) and (**2a**) (pp. 5, 6) for values of n for which n is not the number of vertices of a regular triangular polyhedron. The corresponding packings of spherical caps on the unit sphere and coverings of the unit sphere by such caps are called *irregular*.

A simple consideration of convergence shows that both the extremal problems stated must have (at least) one solution for each integer n. For the packing problem the radius of the cap is in fact bounded from above by II (28). There is therefore an upper bound r_n for the (spherical) radii of n congruent spherical caps which can be packed on the unit sphere without overlapping. Let $r_{n\nu}$ be a subsequence, converging to r_n, from the set of radii of caps corresponding to all possible configurations of n points of the unit sphere. Then there exists a subsequence $r'_{n\nu}$ of $r_{n\nu}$ such that all the sequences of n centres $C'_{N\nu}$ ($N = 1, 2, \ldots, n$) converge, say to $C_1, C_2, \ldots, C_n$. It is now easily seen that the spherical caps of radius r_n drawn about these points C_N ($N = 1, 2, \ldots, n$) can have no interior points in common. Moreover, for reasons of continuity, the systems of caps with centres $C'_{N\nu}$ (and the corresponding radii $r'_{N\nu}$) must have such overlappings from a certain number ν onwards.

Hence it has been shown that there exists a system of caps with extremal radius for the integer n. The existence of a solution for the covering problem can be established in a similar manner.

We may now find the extremal distributions of points on the sphere and determine the corresponding radii r_n for individual integers n. Since the problem has already been solved for the integers $n = 3$, 4, 6 and 12, let us consider first of all the integers 5, 7, 8, 9, 10, 11.

We start with the packing problem and first deal with the case $n = 5$. It is obviously true that

$$r_{n-1} \geqq r_n \geqq r_{n+1} \tag{1}$$

for the limiting radius r_n of the cap corresponding to the number n and therefore from the table on p. 24 we have:—$54° 46' \geqq r_5 \geqq 45°$. We want to show now that $r_5 = r_6 = 45°$.

If three of the five centres of a packing of five non-overlapping caps of (spherical) radius r are joined by great circles, then we have a spherical triangle to which we can apply the cosine formula of spherical trigonometry:

$$\cos c - \cos a \cos b = \sin a \sin b \cos \gamma. \tag{2}$$

Let us now assume that r is greater than 45°. Then the sides a, b and c of the triangle would be greater than 90°, since the caps of radius r must not overlap. Therefore the left, and hence also the right side of (2), would be negative. Since $\sin a$ and $\sin b$ are positive, $\cos \gamma$ would be negative and therefore γ an obtuse angle. However, this would mean that all the angles in all the triangles which one can construct from the 5 centres C_ν ($\nu = 1, 2, 3, 4, 5$) would have to be obtuse. But this is impossible:—If one joins, say, C_1 to the 4 other centres then one would have 4 obtuse angles at C_1. Therefore the supposition $r > 45°$ is wrong. Thus we have $r_5 \leqq 45°$ and, from (1), $r_5 = 45°$.

Thus we have obtained an astonishing result:—The radius of the caps for $n = 5$ is no greater than that for $n = 6$. From this we obtain an extremal packing for $n = 5$ by drawing caps of radius $r_5 = r_6$ round any 5 of the 6 vertices of an octahedron inscribed in the sphere.

2. The Packing Problem Graph

Now suppose we are given a system of n ($n \geqq 7$) points on the unit sphere with minimum spherical distance apart $2r$. We introduce the concept of the *graph* in order to determine the conditions under which such a system of points represents a solution of our problem.

The graph is a system of points and lines (arcs of great circles) on the sphere. The following belong to it:—

1. The n points of our system C_ν ($\nu = 1, 2, 3, \ldots, n$),
2. all lines $C_\nu C_\mu$ for which the (spherical) length $C_\nu C_\mu = 2r$.

First of all we note that the lines of a graph cannot intersect each other in interior points. Obviously, if the lines $C_\nu C_{\nu+1}$ and $C_\mu C_{\mu+1}$ had a point S in common, then the following would be true:—

$$4r \leqq C_\nu C_{\mu+1} + C_\mu C_{\nu+1} < C_\nu S + SC_{\mu+1} + C_\mu S + SC_{\nu+1}$$
$$= C_\nu C_{\nu+1} + C_\mu C_{\mu+1} = 4r.$$

But this is a contradiction.

If λ lines intersect at a point C_ν, then we say that the point is *of order*, λ. Isolated points are then of order 0.

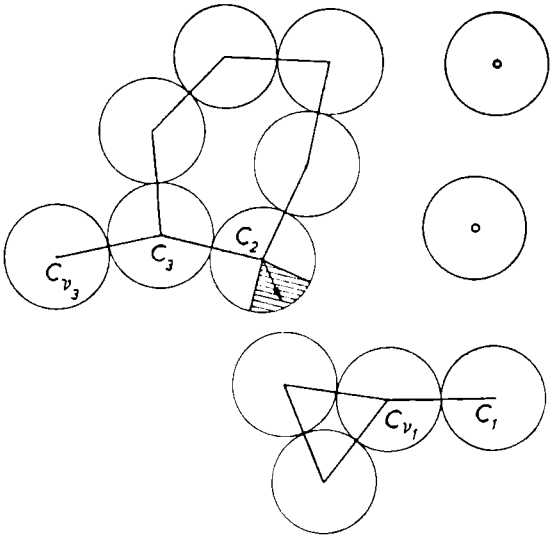


FIG. 9

If a graph contains points of order 1 or order 2 then, by displacing individual points on the sphere, one can get a new graph which has *fewer* lines (for the same r). Let there be a point, say C_1 (fig. 9), of order 1 which is joined by a line of the graph to *one* point C_{ν_1}. Then one can displace C_1 with its cap a small distance in the opposite direction to $C_1 C_{\nu_1}$ without the cap overlapping any other cap of the system. Obviously the same is possible for points of order 2 (C_2 in fig. 9):—In this case we determine the admissible direction for the displacement by the angle enclosed by the perpendiculars at C_2 to the two lines of the graph. One can carry out a small displacement of C_2 (with cap!) within this sector without the " misfortune " of overlapping another cap. Such a displacement is not possible for the point C_3 (of order 3) in fig. 9 since the cap with centre C_{ν_3} makes a displacement in the direction of displacement permitted by the two other caps impossible.

If this process of displacement can be continued until *all* the points are isolated then we have obtained from the original point system $\mathcal{S}$ a new system $\mathcal{S}'$ also consisting of n points but in which the least distance between any two points is *greater than* $2r$. This means that the given system $\mathcal{S}$ is not a configuration having maximum density of packing.

We now introduce the term an *irreducible* graph for a connected system of lines of a graph in which no further displacements are possible. Therefore to solve the extremal problem in question we have to investigate point systems with irreducible graphs. For these investigations we need a †

Lemma: *If an irreducible graph contains a polygon with* μ *vertices then* $2\mu r < 360°$.

To prove this we form a spherical triangle with the same perimeter as the polygon. This is done most simply by imagining our polygon to have rods as sides, these being circular arcs of length $2r$ and curvature 1. We can now displace this polygon of rods so that we obtain from it a triangle whose sides are multiples of $2r$. Its perimeter is the same as that of the polygon, namely $2r\mu$. However, since the perimeter of every spherical triangle is smaller than that of a lune which is not convex, it is therefore smaller than 360°; and hence the lemma has been proved.

In particular, for $\mu = 5$ we have $2r < 72°$. If it is established now for a number n that there is an irreducible graph with $2r > 72°$ then it is also established that there cannot exist a pentagon in this graph and, even more so, in the irreducible graph with greatest radius r.

3. A Point System for $n = 7$

To solve our extremal problem for $n = 7$, we mention first of all a particular configuration of 7 points on the unit sphere for which $2r > 77°$. We shall see later that this configuration already represents the solution of our problem. Hence it is expedient to show first of all the existence of a particular solution with $2r > 77°$. For we obtain from it an important result:—From the corollary just deduced from the lemma, there cannot be any polygons with more than 4 vertices in the graph of an extremal configuration (for which, of course, $2r \geqq 77° > 72°$).

† In the following we measure the angles *and hence the surface areas of spherical polygons in degrees.*

To draw the 7-point graph we start from one of the regular tetrahedra inscribed in the unit sphere. The projection of the tetrahedron from the centre of the sphere yields a partitioning of the unit sphere into 4 congruent, equilateral spherical triangles. We call the centre of one of these 4 triangles the south pole of the unit sphere.

Now we replace the tetrahedron by another figure:—We start again from an equilateral triangle ABC whose centre must lie at the south pole. However, its sides $2r$ must be smaller than the sides

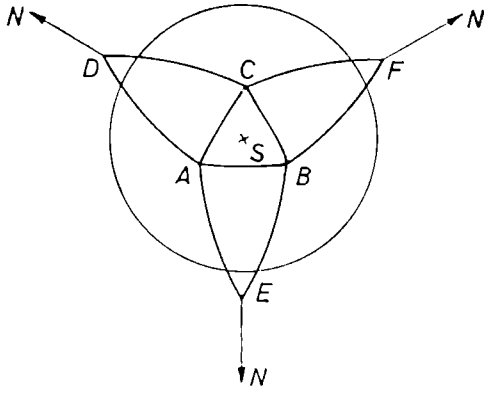


FIG. 10

$2r_4$ of the tetrahedron. At each of the three sides of the triangle we place a congruent, equilateral triangle: ABE, BCF and CAD. Then the three points D, E and F do not coincide at the north pole of the sphere this time, but they have the same geographical latitude. If one chooses the base $2r$ small enough then the lines DN, EN and FN will be greater than $2r$. Since the lines were $DN = EN = FN = 0$ in the case of the tetrahedron, between $2r = 0$ and $2r = 2r_4$ there must be a value of r for which the lines DN, EN and FN have length $2r$ exactly. The 7 points A, B, C, D, E, F and N obtained in this way form the special point system $\mathscr{S}_7$ which we now want to consider. Fig. 10 shows $\mathscr{S}_7$ in stereographic projection.†

We have joined the 7 points by lines so that the whole sphere is partitioned into 4 equilateral triangles and 3 rhombuses. Also, since the opposite angles are equal in a spherical rhombus, we have

$$\angle DNE = \angle ENF = \angle FND = \angle EAD = \angle FBE = \angle DCF = 120°.$$

† In connection with this see Chapter X, section 1.

If we denote the angle of the 4 equilateral triangles by α, then it follows for the sum of the angles at A, B or C that $3\alpha+120^\circ = 360^\circ$, therefore $\alpha = 80^\circ$. From this we calculate $2r$:—In the half equilateral triangle with side of length $2r$ and angle $\alpha = 80^\circ$, we have by the rules of spherical trigonometry

$$\sin\frac{\alpha}{2} = \frac{\sin r}{\sin 2r} = \frac{1}{2\cos r}, \quad 2r \sim 77^\circ\, 52'\, 10''.$$

Therefore for the (still unknown) extremal radius of the cap for $n = 7$ it is true that

$$2r_7 \geqq 2r \sim 77^\circ\, 52'\, 10''.$$

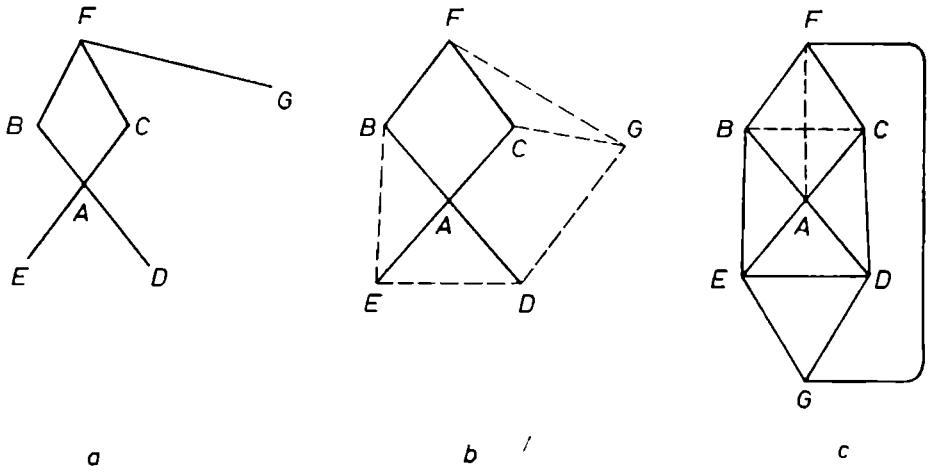
Therefore, by the lemma on p. 32 the only polygons in the extremal graph are triangles and quadrangles. Now we still want to add that there cannot be any isolated points of the graph in the interiors of the triangles and quadrangles. To see this, we have only to observe that every point in the interior of an equilateral spherical triangle or quadrangle (with side $2r < 90^\circ$) is at a distance of less than $2r$ from the vertices. For the triangle this result is trivial. For a rhombus $KLMN$ with $KL = 2r < 90^\circ$ the circles with radius $2r$ about K and M are *not great circles*: therefore they *overlap* the whole rhombus and every point in the interior of the rhombus is at a distance of less than $2r$ from the vertices. Such a point cannot belong to the graph of a packing system of caps of radius r. Hence, for the case $n = 7$, we may assume that the extremal graph contains no isolated points and consists only of triangles and quadrangles.

4. The Packing Problem for $n = 7$

Having completed the preparatory work we can turn to the solution of our extremal problem for $n = 7$. The required graph contains only triangles and quadrangles: it has no isolated points. Not all of its points can be of order 3, otherwise we would have for the number of edges e the relation $3.7 = 21 = 2e$. This is impossible since e must be integral. Therefore there must be at least one of the 7 given points which is of higher order. However, points of order 5 cannot appear in an extremal 7-point graph. Of 5 lines starting from a point at least two would enclose an angle less than 72°. However, in our example given in the preceding section, the angle of the equilateral triangle was already 80° in magnitude. Therefore the equilateral triangle in an extremal graph must have

an angle which is *at least* equal to 80°. An isosceles triangle with equal sides of length $2r$ and an angle at the apex which is at most equal to 72°, would therefore have base *less than* $2r$. Hence we would have two points in the graph which would intefere with each other. Therefore we exclude points of order 5 (and of course points of higher order).

Now let A be a point of an extremal graph of order 4. The 4 lines of the graph starting from A may lead to the points B, C, D and E (fig. 11a).



FIGS. 11a-11c

The graph contains two other points F and G. Since they are at least of order 3, F and G are joined to at least two of the points B, C, D and E. However, these points cannot be opposite each other. Obviously, if one joined, say, F with B and D, then (since intersections are impossible) C (or E) would be a point of order 1. However, such points do not exist in an irreducible graph. Therefore let F be joined to B and C (fig. 11a). Then, from F (being a point of at least third order) another line leads to the 7th point G. G must then be joined to at least two of the 4 points B, C, D and E, which are not opposite each other. Therefore, it would be possible to take C and D, D and E, or E and B.

Let us consider the first case. Since B, E and D must be at least of order 3, the lines BE and ED would also have to be drawn into the graph (fig. 11b). However, in this way, we obtain a graph with a pentagon ($BEDGF$), which is not divided by diagonals belonging to the graph. Since the lines of the extremal graph are certainly not

greater than 90°, the diagonals FE, BD etc., lying *in the interior* of the pentagon, have length greater than $2r$. Since a pentagon cannot appear in the graph, this case is excluded. This is also true for the pair of points E and B, hence G must be joined to the points D and E (fig. 11c). Since only triangles and quadrangles may appear, we must also include the lines BE and CD.

This graph has still one degree of freedom: the angle BAC is not yet determined. Certainly it is clear that it cannot be smaller than that of the equilateral triangles of the graph. Otherwise the points B and C would interfere with each other: their distance apart would be smaller than $2r$.

Let us denote by β the (still undetermined) angle BAC and by α_7 the angle of the equilateral triangle. Then it is true that

$$\alpha_7 \leqq \beta \leqq 360° - 3\alpha_7, \tag{3}$$

since the angle $EAD = \beta_1 = 360° - 2\alpha_7 - \beta$ also may not be less than α_7. Now, if one divides the rhombus $ACFB$ by its diagonals into four right-angled, spherical triangles, one obtains from these triangles the relation

$$\tan \frac{d}{2} = \tan 2r \,.\, \cos \frac{\beta}{2}$$

for the diagonal $d = AF$. Now, a simple calculation shows that the second derivative of

$$d = f(\beta) = 2 \tan^{-1} \left(\tan 2r \,.\, \cos \frac{\beta}{2} \right)$$

is negative and therefore the function $f(\beta)$ is concave (by the criterion on p. 13). Likewise $d_1 = AG = f(\beta_1)$ is a concave function of β_1. But

$$d_1 = f_1(\beta) = f(360° - 2\alpha_7 - \beta)$$

is also a concave function of β, as one can see by the same criterion. Since the sum of two concave functions is also concave,† $FG = d + d_1 = d_2$ is a concave function of β. Because a function which is concave in an interval $[a, b]$ does not assume its minimum in the interior of the interval, $FG = d_2(\beta)$ does not become the minimum for a value of β in the interior of the interval (3).

Let us now assume that we are given an extremal graph for $n = 7$ with an angle β from the interior of (3). Let us imagine the lines of

† This follows from the definition II (5').

this graph, with the exception of the line FG, to be formed in practice by a system of rods on the unit sphere. By a small change of the angle $\beta = \angle BAC$ one can now change this system of rods so that $d_2 = FA + AG$ becomes smaller and therefore the *complementary* line FG (belonging to the graph up till now) becomes larger. This is possible because, as was just established, $d_2(\beta)$ does not have a minimum in the interior of (3). The graph is disrupted by this displacement, for the line FG is now greater than $2r$. However, one can now construct a new system of lines with a greater side $2r_1$ *retaining the angle* β. From continuity considerations there must be a value $r_1 > r$ for which the complementary line FG is just equal to $2r_1$. Then, however, one has again obtained a graph as prescribed with side $2r_1 > 2r$. But this means that the original graph given was not extremal.

Hence it follows that we obtain an extremal graph for $n = 7$ only if we put the angle β (or β_1 respectively) equal to α_7 in one of the two rhombuses. Then this rhombus consists of two equilateral triangles and hence the diagonal (BC or ED respectively) of the rhombus must still be drawn to make the graph complete:—This diagonal joins the points which are exactly at a distance $2r$ apart.

If, in addition, in fig. 11c we plot the diagonal BC, we obtain just the schematic representation of the graph already represented in stereographic projection in fig. 10. In fact:—The equilateral triangle ABC is adjoined in both cases by three congruent, equilateral triangles ABE, BCF and CAD. The point denoted in fig. 10 by N is called G in fig. 11c, otherwise the notation is the same. Hence it has been shown that the graph introduced in section III 3 solves our extremal problem.

We have now solved the problem (**1a**) and hence also problem (**1**) of p. 1, for a series of numbers n and we want to summarise the result:—

Theorem 11: *The problem of arranging n points on the unit sphere so that the distance between any two of the points is as great as possible is solved for the numbers* $n = 4$, 6 *and* 12 *by the vertices of the regular, inscribed, triangular polyhedra (tetrahedron, octahedron, icosahedron). For* 5 *points one obtains the maximum for the minimal distance apart if one places them at* 5 *of the* 6 *vertices of an octahedron. For* $n = 7$, *the extremal graph consists of four equilateral triangles (with angle* $\alpha_7 = 80°$) *and three rhombuses (fig.* 10).

We can describe the solution discovered by SCHÜTTE and VAN DER WAERDEN [III 1] for $n = 7$ in the language of mathematical geography as follows:—

> *One puts an equilateral triangle with angle* $\alpha_7 = 80°$ *on the sphere so that its centroid coincides with the south pole and adjoins to each of the three sides of the triangle another congruent triangle. Then the six vertices of these four triangles together with the north pole form a 7-point configuration which solves the problems* **(1)** *and* **(1a)** *for* $n = 7$.

To discover the extremal graph for larger values of n becomes ever more complicated. The solutions for $n = 8$ and $n = 9$ have been stated by SCHÜTTE and VAN DER WAERDEN. There exist conjectures ([III 1] and [II 7]) for other numbers.

Later, we shall collect the results known up to the present time in a table (p. 48). We now turn our attention to problems **(2)** and **(2a)**.

5. THE COVERING PROBLEM GRAPH

The covering problem (problems **(2b)** and **(1b)** of p. 6) has been solved up till now only for numbers n which are the numbers of vertices for regular triangular polyhedra (corollary of Theorem 9). To find the solution for any number n, we start from an arbitrary covering of the unit sphere by n congruent spherical caps (of spherical radius r). If every point of the sphere is covered by at least one *interior* point of the system of caps, it is obviously possible to replace the system of caps by another, with the same centres but somewhat smaller spherical radius $r'(r' < r)$, which is also a complete covering of the sphere. Then the original system given was not one of minimal density of packing. Therefore, if we are looking for a covering of the sphere by n caps with spherical radius as small as possible, we can limit ourselves to those cases for which certain points of the sphere are covered only by (3 or more) boundary circles of the caps. In the following we shall call such points of intersection " I-points ". For the centres of the caps we introduce the abbreviated notation " C-points ".†

Now we join (fig. 12) the I-points of a covering of the sphere to the pertinent C-points by (spherical) lines of length r. In this way a " trihedral " (see I_1 in fig. 12) or a " tetrahedral " (I_2), " penta-

† In Fig. 12 and the other figures of this chapter we indicate the I-points by small circles and the C-points by large dots.

hedral " etc. is associated with each I-point. The configuration of all such lines of length r is called a graph associated with the spherical covering.† If there exist several trihedrals (m-hedrals) in a covering then it is possible that they combine together to form polygons. The number of their vertices is always even since C- and I-points alternate in the polygon (fig. 12).

Here also we can try to reduce the graph by a suitable transformation of caps and if possible reduce it completely. If this succeeds

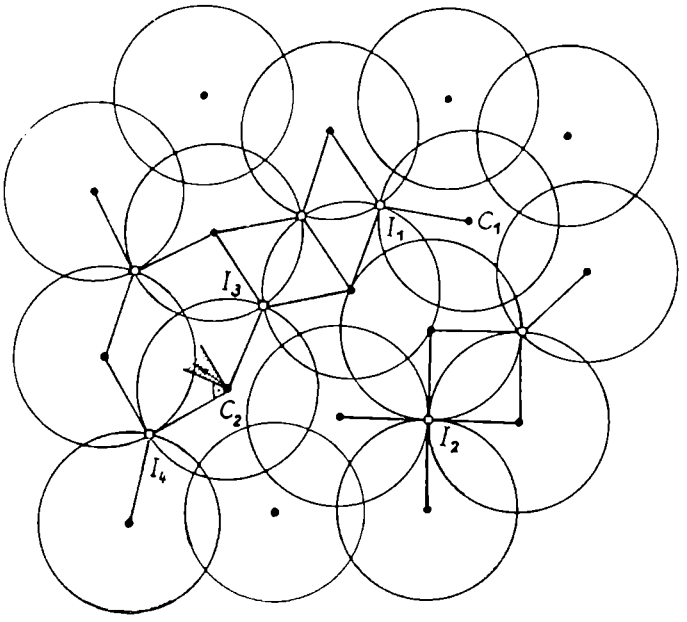


FIG. 12

then the given covering and the one obtained from it by displacement are certainly not extremal. In this case our problem will again be to discover coverings which have an irreducible graph.

Trihedrals which at all three C-points do not adjoin another trihedral (m-hedral), can be eliminated by a displacement. If one displaces for example the cap with centre C_1 in fig. 12 in the direction of the I-point I_1, we obtain at I_1 a small spherical triangle which is covered three times over. The covering of the sphere is not disturbed at any other place for a sufficiently small displacement. In this way we obtain a new covering which has the same density of packing and for which the graph has become smaller: the trihedral

† This definition of the graph differs from that which SCHÜTTE introduced in his work [III 2] on this problem.

at I_1 must be cancelled since I_1 is not an I-point for the new covering. It is also possible to reduce the graph by a displacement if a C-point of order 2 appears in a graph, e.g. C_2 in fig. 12. The two I-points to which C_2 is joined in the graph, are indicated in fig. 12 by I_3 and I_4. Now, if one displaces the cap with centre C_2 into the interior of the angle formed by the perpendiculars to C_2I_3 and C_2I_4 at C_2 then for a sufficiently small displacement the complete covering of the sphere is not disturbed but the points I_3 and I_4 are no longer I-points. Again this causes a reduction of the graph.

From all this we see that an *irreducible* (therefore no reduction possible) graph has the following properties:—

1. It is made up of lines of length r.
2. The lines form polygonal nets with even numbers of vertices.
3. The vertices of the graph are at least of order 3.
4. The lines of the graph do not intersect each other.

One can prove the last property as follows: If two lines $C_\nu I_\nu$ and $C_\mu I_\mu$ had a point of intersection T then we would have

$$C_\nu I_\mu + C_\mu I_\nu < C_\nu T + TI_\mu + C_\mu T + TI_\nu = C_\nu I_\nu + C_\mu I_\mu = 2r.$$

But this means that at least one of the lines $C_\nu I_\mu$ and $C_\mu I_\nu$ is smaller than r. Then I_ν (or I_μ) would be covered by the circle about C_μ (C_ν) and would not be an I-point.

6. The Covering Problem for $n = 5$

Now let us find the graph for an extremal covering of the sphere by 5 congruent caps. It is obvious that a hexagon cannot appear in a graph with 5 C-points. All the polygons must be rhombuses and all C-points are of order 3 or 4. If a C-point C_1 of order 4 appears then already all 5 C-points appear in the 4 rhombuses meeting at C_1 (fig. 13—a schematic diagram).

Then, from the 4 opposite vertices C_2, C_3, C_4, C_5 of the first C-point C_1, there are only lines leading to an I-point we can imagine lying at the north pole of the sphere. Hence the sphere is partitioned into 8 rhombuses.

However, a C-point of order 4 need not necessarily occur in the graph. If we start from a C-point C_1 where three rhombuses meet (fig. 14)—$C_1I_1C_2I_2$, $C_1I_1C_4I_3$ and $C_1I_3C_3I_2$—, we can adjoin to this " star " three further rhombuses (with I-points I_4, I_5 and I_6)

without using new C-points. Since only one C-point remains at our disposal, the lines from I_4, I_5 and I_6 must lead to the fifth C-point which we again imagine at the north pole.

First of all we deal with the symmetrical graph of the type represented in fig. 14. Let

$$\angle I_1C_1I_2 = \angle I_1C_1I_3 = \ldots = \angle I_4C_5I_5 = \angle I_5C_5I_6 = 120°.$$

Then the 6 rhombuses meeting at C_1 and C_5 are congruent and their diagonals $C_1C_2 = C_2C_5 = \ldots = 90°$ in magnitude. The three rhombuses meeting at C_1 are each divided by their diagonals into 4 congruent right-angled triangles with hypotenuse r, an angle

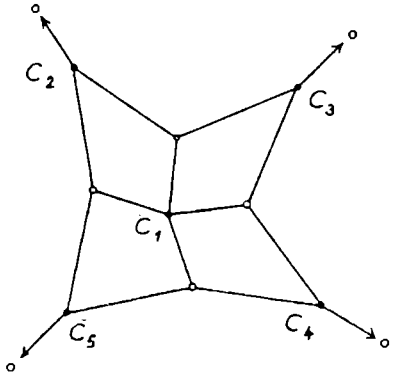
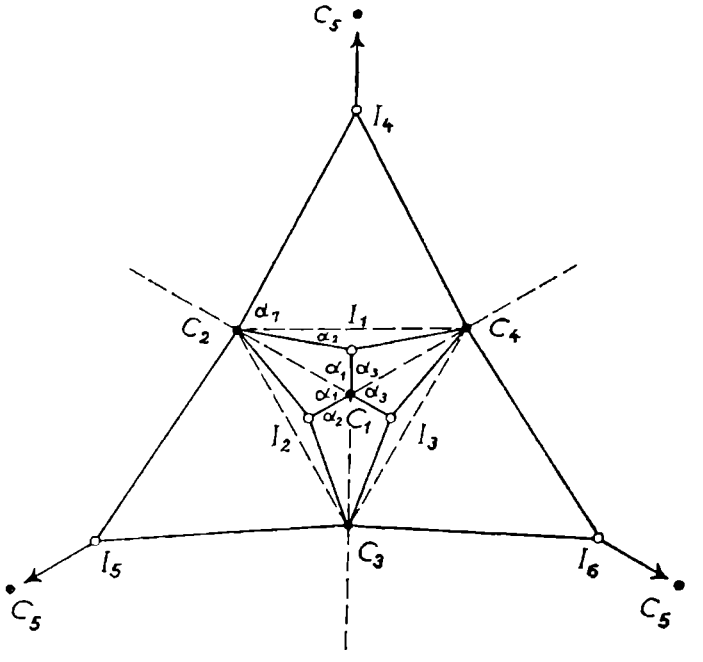


FIG. 13 (left)

FIG. 14 (below)



of 60° and the other arm of this angle 45°. Hence one determines at once

$$\tan r = \frac{1}{\cos 60°} = 2,\ r \sim 63° 26'. \tag{4}$$

In the general (non-symmetrical) case we have the rhombuses with angles $2\alpha_\nu$ ($\nu = 1, 2, 3$; $\nu = 4, 5, 6$) meeting 3 at a time at C_1 and C_5. Further, let

$$\angle I_\nu C_{\nu+1} I_{3+\nu} = 2\alpha_{6+\nu},\quad \nu = 1, 2, 3.$$

Finally, we denote the other angles of these 9 rhombuses by $2\beta_\nu$ ($\nu = 1, 2, \ldots, 9$). Then the following are true:

$$2\sum_{\nu=1}^{3} \alpha_\nu = 2\sum_{\nu=4}^{6} \alpha_\nu = 360° \tag{5}$$

and

$$2\sum_{\nu=1}^{6} \alpha_\nu + 4\sum_{\nu=7}^{9} \alpha_\nu = 3 \,.\, 360°,$$

therefore

$$2\sum_{\nu=7}^{9} \alpha_\nu = 180°. \tag{6}$$

The 9 rhombuses of the graph are now decomposed into 18 triangles by the diagonals joining the C-points. Since

$$\cot \alpha_\nu \cdot \cot \beta_\nu = \cos r \quad (\nu = 1, 2, \ldots, 9),$$

we can write the surface area of such a triangle as follows:

$$a(\alpha_\nu) = 2(\alpha_\nu + \cot^{-1}(\cos r \,.\, \tan \alpha_\nu)) - \pi.$$

Hence it follows that

$$4\pi = -18\pi + 4\sum_{\nu=1}^{9} (\alpha_\nu + \cot^{-1}(\cos r \,.\, \tan \alpha_\nu)). \tag{7}$$

Now the second derivative of $a(\alpha_\nu)$ is negative for $\alpha_\nu < 90°$, $r < 90°$:

$$a''(\alpha_\nu) = \frac{4 \cos r(\cos^2 r - 1) \sin 2\alpha_\nu}{(\cos^2 \alpha_\nu + \sin^2 \alpha_\nu \cos^2 r)^2} < 0.$$

Hence the function $a(\alpha_\nu)$ is concave and we can apply Jensen's theorem (Theorem 3) to each of the sums from 1 to 3, from 4 to 6 and from 7 to 9 in (7). Then, by (5) and (6), it follows from (7) that

$$22\pi \leqq 10\pi + 24 \cot^{-1}\left(\cos r \,.\, \tan \frac{\pi}{3}\right) + 12 \cot^{-1}\left(\cos r \,.\, \tan \frac{\pi}{6}\right).$$

therefore

$$2\left(\frac{\pi}{2}-\cot^{-1}(\cos r\,.\,\sqrt{3})\right)\leqq\cot^{-1}(\cos r\,.\,\tfrac{1}{3}\sqrt{3}). \tag{8}$$

By the substitutions

$$\tan\phi=\cos r\,.\,\sqrt{3},\ \cot\psi=\cos r\,.\,\tfrac{1}{3}\sqrt{3},\ \tan\psi=u \tag{9}$$

from (8) we have

$$2\phi\leqq\psi,$$

$$\tan 2\phi=\frac{2\tan\phi}{1-\tan^2\phi}=\frac{2\,.\,3}{u\left(1-\dfrac{9}{u^2}\right)}\leqq\tan\psi=u.$$

Therefore we have

$$u=\tan\psi\geqq\sqrt{15}$$

and from (9)

$$r\geqq\cos^{-1}\tfrac{1}{5}\sqrt{5}=\tan^{-1}2.$$

But this means that for every graph of the type represented in fig. 14, r is not less than in the symmetrical case for which, by (4), $\tan r=2$ is valid.

The graphs of the type represented in fig. 13 (with a C-point of order 4) also do not yield a smaller value for r. One quickly calculates that in the symmetrical case

$$\cos r=\cot 45^\circ\cot 67{\cdot}5^\circ,$$

therefore $r\sim 65^\circ\,32'$. For the graph of this type, applying Jensen's theorem also leads to the statement that it is not possible to have a cap radius less than that of the symmetric type. Therefore the type of symmetric graph represented in fig. 15 in stereographic projection *with no* C-points of order 4 yields the best possible value † for r for 5 C-points:—

$$r=R_5=\cos^{-1}\tfrac{1}{5}\sqrt{5}=\tan^{-1}2\sim 63^\circ\,26'.$$

In geographical terminology one can describe this graph simply as follows:—If there is one point at the north pole and one at the south pole then the three others are distributed symmetrically on the equator. In the case $n=6$ we have the corresponding solution:— In this case the extremal configuration consists of the vertices of an

† We will denote by R_n the extremal value of r for the *covering* by n caps. We have used the notation r_n for the packing problem.

inscribed octahedron and, of course, one can also place this body so that one point is at the north pole and one at the south pole and the others lie on the equator.

Therefore it is natural to conjecture that for more than 6 points the extremal configuration may also be formed by a corresponding distribution. In any case it is not meaningless to calculate the graph of this type for any n.

The $(n-2)$ rhombuses meeting in the north or the south pole respectively are each divided by their diagonals into 4 congruent,

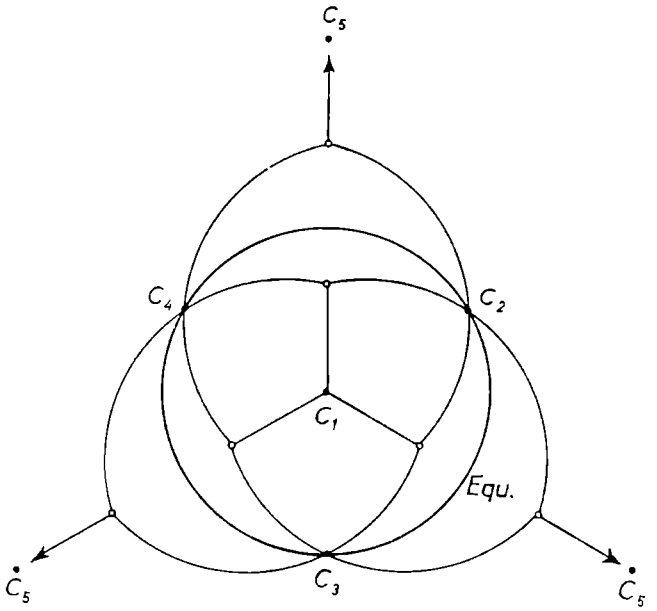


FIG. 15

right-angled triangles which have an angle (at the north or south pole respectively) of $\left(\frac{180}{n-2}\right)^\circ$ and an arm (of this angle) of 45°. The side r for $n = 7$ is calculated from this triangle:—

$$r = \rho_7 = \cot^{-1}(\cos 36^\circ) \sim 51^\circ\, 2'.$$

In the case $n = 8$

$$\rho_8 = \cot^{-1}(\cos 30^\circ) \sim 49^\circ\, 7'. \tag{10}$$

It remains to be shown whether or not these values ρ_7 and ρ_8 are equal to the required extremal values R_7 and R_8.

7. The Covering Problem for $n = 7$

To survey the graphs coming into question in the case of the covering problem for $n = 7$, we introduce a simpler system of lines and call it the "C-figure". It corresponds to the graph † used in the packing problem. If, in a graph made up of rhombuses, one draws the diagonals joining the C-points in each rhombus then, on cancelling the graph, there results a triangular net joining the centres which is in

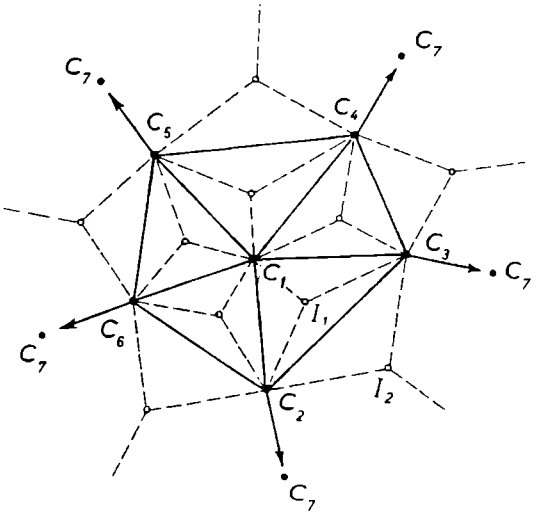


Fig. 16

fact the C-figure. The C-figure is indicated in fig. 14 by the dashed lines.

If, conversely, a C-figure is given, one can recognise the graph, at least in its topological structure:—One plots in each triangle an I-point and joins it to the vertices of the triangle (fig. 16). On cancelling the C-figure one obtains the graph. At the same time one must take into account that I-points of order 4 might arise;‡ in this case a rhombus would degenerate into two lines, each traversed twice, and I-points would concide e.g. I_1 and I_2, say, in fig. 16.

† In Schütte [III 2] this figure is called the graph.

‡ One can certainly show easily that in the case $n = 7$ no I-points of order 4 can arise. Four cap circles which pass through the south pole, say, and are of radius $R_7 \leqq \rho_7 \sim 51° 2'$ would leave uncovered a cap of radius 77° about the north pole. Now the following theorem, whose proof we leave to the reader, is valid:—*If a cap of spherical radius ρ is covered by three congruent caps of radius r, then*

$$\rho \leqq 2 \tan^{-1} (\tfrac{1}{2} \tan r).$$

Hence it follows that the uncovered cap about the north pole cannot be covered by the three remaining caps of radius R_7.

From Euler's theorem (Theorem 2) for a polyhedron and the relation $3a = 2e$ one immediately obtains for $n = 7$ the number a:— The C-figure consists of 10 triangles. We assert that none of the C-points can be of order 3. If there were,† then obviously one could combine three of the 10 triangles into *one* triangle. Since the sides of this triangle are diagonals of the graph, their length is less than $2r$. The area of this triangle is therefore at most equal to the area T of an equilateral triangle with sides of length $2r$. However, the area of the 7 remaining triangles is at most equal to the area $\mathscr{T}$ of an equilateral triangle inscribed in the circle of radius r. For the angle α_1 of the first triangle the following is true

$$\cos r \,.\, \sin \frac{\alpha_1}{2} = \cos 60°;$$

on the other hand we have for the angle of the second triangle

$$\cot \frac{\alpha_2}{2} \,.\, \cot 60° = \cos r.$$

Hence, it follows when one translates the radian measure into degrees for $r \leqq \rho_7 < 52°$ that

$$\mathscr{T} = 6 \cot^{-1} (\sqrt{3} \cos r) - 180° < 80°$$

and

$$T = 6 \sin^{-1} \left(\frac{1}{2 \cos r} \right) - 180° < 150°.$$

Since the 10 triangles partition the whole sphere, it follows that

$$720° \leqq T + 7\mathscr{T} < 710°.$$

Because of this contradiction, none of the C-points can be of order 3. Since, obviously, no C-points of order 6 can arise, we have, if c_ν denotes the number of C-points of order ν,

$$4c_4 + 5c_5 = 2e = 30,$$

$$c_4 + c_5 = n = 7.$$

Therefore $c_5 = 2$, $c_4 = 5$. But a C-figure with these numbers leads to the type represented in fig. 16. Hence, in the case of symmetry, one obtains the graph whose side ρ_7 we have already calculated.

† Here we are following SCHÜTTE [III 2].

Here two of the C-points are at the north and south pole of the sphere and the other five are distributed at equal distances apart on the equator. On applying Jensen's theorem (Theorem 3), one shows again, as in the case $n = 5$, that no unsymmetrical ordering of C-points leads to a graph of this type having a smaller value for r.

Hence it has been shown that $R_7 = \rho_7 = \cot^{-1}(\cos 36°)$ is the smallest value for r in a graph consisting of *only rhombuses*. We leave the reader to prove that in the case $n = 7$ the extremal graph cannot contain a hexagon and we summarise our results in

Theorem 12: *One obtains a covering of the unit sphere by* 5, 6 *or* 7 *congruent spherical caps of smallest possible radius* R_n $(n = 5, 6, 7)$ *when two of the centres of the caps are placed at the north and south poles of the sphere and the others are distributed at equal distances apart on the equator. For the smallest possible radius* R_n *of the spherical cap*

$$R_n = \cot^{-1}\left(\cos\frac{\pi}{n-2}\right)$$

is true.

8. An Outline of Further Problems

Anyone who imagines that this can be extended similarly is wrong. For $n = 8$ there is [III 2] a graph whose radius r is smaller than the value (10), which was calculated in Section III 6 and which is the

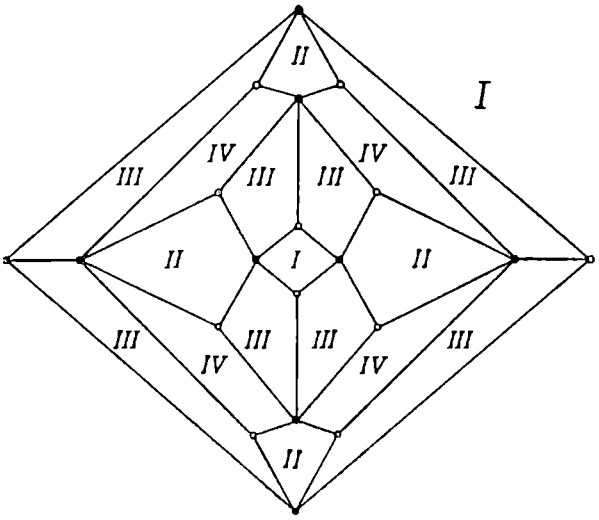


Fig. 17

value for the equatorial distribution of the points. This graph, represented schematically in fig. 17, is not so easy to deal with, even in the symmetrical case. One is led to a transcendental equation with a solution

$$r \sim 48^\circ\, 9' < \rho_8 \sim 49^\circ\, 7'.$$

By an argument similar to that for $n = 7$ one can show, using this result, that there cannot be a C-point of order 3 in the C-figure for $n = 8$. Hence one can deduce that only a graph of the type represented in fig. 17 must be considered for our extremal problem when $n = 8$ [III 2].

Let us collect the most important results of Chapters II and III in a table. The first two rows give the extremal radius and the density of packing *for the packing problem*, the following rows refer to the covering problem. We have noted not only the radius and packing density (R_n and D_n) but also the number of the C- and I-points of order v (c_v and i_v respectively).

	4	5	6	7	8	9	10	11	12
r_n	54° 46′	45°	45°	38° 56′	37° 26′	35° 16′			31° 43′
d_n	0·845	0·732	0·878	0·777	0·823	0·825			0·896
R_n	70° 32′	63° 26′	54° 44′	51° 2′	$\leqq$48° 9′				37° 43′
D_n	1·33...	1·382	1·268	1·299	$\leqq$1·331				1·254
c_3	4	2	0	0	0				0
c_4	0	3	6	5	4				0
c_5	0	0	0	2	4				12
i_3	4	6	8	10	12				20

For some of the still unsolved cases of the packing problem ($n = 10, 11, 13, \ldots$) there exist approximations and conjectures, discussed in FEJES TOTH [II 7]. Perhaps it is possible in the cases which are still unsolved to make some progress by looking for regularities in the structure of the graphs. We have also recorded in the table the numbers c_v and i_v for the covering problem. i_4 is not noted since an I-point of order 4 did not occur in any of the extremal graphs found so far. However, at present, it has not been proved that this must always be the case. It is natural to interpolate between the values determined for $n = 8$ and $n = 12$ in the rows for c_3, c_4, c_5 and i_3 as follows:

	9	10	11
c_3	0	0	0
c_4	3	2	1
c_5	6	8	10
i_3	14	16	18

Chapter IV

PACKING OF CONGRUENT SPHERES

1. The Densest Lattice Packing

So far in dealing with packing problems it has always turned out that the extremal configurations exhibit a high degree of symmetry. If one *assumes* certain symmetry properties then finding the solution becomes intrinsically easier: of course, the results obtained cannot be formulated as generally as for the procedures " without assumptions ", which we used in Chapters II and III.

To make this difference clear let us consider once again the problem of packing unit circles in the plane. If we assume that the centres of the circles have been arranged to form a lattice of parallelograms, we can very quickly see how to get the densest packing: one must choose the parallelograms to be rhombuses with angles of 60° and 120° (fig. 4).

Let a point lattice like this be given by the position vectors

$$\mathfrak{z} = m\mathfrak{a} + n\mathfrak{v} \tag{1}$$

starting from some zero, where $\mathfrak{a}$ and $\mathfrak{v}$ are (linearly independent) vectors, whose length is at least equal to 2 and m and n are arbitrary integers. If $|\mathfrak{a}|$ or $|\mathfrak{v}|$ is greater than 2 the density of packing can obviously be improved. Therefore one can assume $|\mathfrak{a}| = |\mathfrak{v}| = 2$. Then the area of a single parallelogram in the net (1) is $a(\alpha) = 4 \sin \alpha = 4 \sin (\mathfrak{a}, \mathfrak{v})$ and for the density of packing $d(\alpha)$ we obtain † from equation II (2).

$$d(\alpha) = \lim_{R \to \infty} \frac{\pi n(R)}{\pi R^2} = \lim_{R \to \infty} \frac{\pi n(R)}{n(R)\, 4 \sin \alpha} = \frac{\pi}{4 \sin \alpha}.$$

The maximum of $d(\alpha)$ occurs when $\sin \alpha$ is a minimum. $\alpha = 60°$ is the condition for the latter. Obviously a smaller angle is not

† One may replace πR^2 under the limit sign by $n(R) \,.\, a(\alpha)$. This is shown by an argument like the one on p. 15.

possible for, if it were, the circles belonging to a pair of opposite vertices would intersect each other. Hence it follows that $d \leqq \frac{\pi}{\sqrt{12}}$ as in equation II (17).

Since the—as yet unsolved—general packing problem (3) (p. 2) for congruent spheres in space is quite difficult, let us begin by examining packings of congruent spheres in space when these are formed on a lattice. Therefore we start from a lattice (fig. 18)

$$\mathfrak{r} = l\mathfrak{u} + m\mathfrak{v} + n\mathfrak{w} \tag{2}$$

where $\mathfrak{r}$, $\mathfrak{v}$ and $\mathfrak{w}$ are linearly independent vectors whose length is at

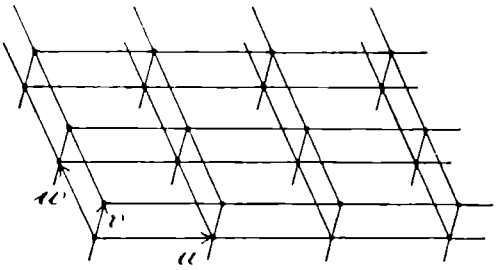


FIG. 18

least equal to 2 and l, m and n are integers. Let the points of (2) be the centres of a system $\{\mathfrak{o}\}$ of unit spheres. Since, here too, one can obviously improve the density of packing by moving all the spheres if one of the vectors is longer than 2, we can limit ourselves to the case where the parallelepipeds formed by the lattice points have only rhombuses as side faces: $|\mathfrak{u}| = |\mathfrak{v}| = |\mathfrak{w}| = 2$.

Now, if $\alpha \neq 60°$, $\neq 120°$ were true, one could, as indicated by the investigation of the density of packing of the unit circles, improve the spherical packing by changing the angle α. The same is true for the $(\mathfrak{u}, \mathfrak{w})$- and the $(\mathfrak{v}, \mathfrak{w})$-planes. Therefore the angles of all the rhombuses of the lattice (2) must be equal to 60° or 120° in the lattice with extremal density.

By a discussion similar to that on p. 15, one obtains for the density of packing $\delta(\mathfrak{o})$ of our system of spheres

$$\delta(\mathfrak{o}) = \lim_{R\to\infty} \frac{N(R)\,\frac{4}{3}\pi}{\frac{4}{3}\pi R^3} = \lim_{R\to\infty} \frac{N(R)\,\frac{4}{3}\pi}{N(R)\,.\,V} = \frac{4\pi}{3V},$$

where V is the volume of the parallelepiped determined by the vectors u, v and w and $N(R)$ is the number of parallelepipeds lying completely within the sphere of radius R. However, for the extremal system of spheres $\{\sigma^*\}$ the volume is $V = V^* = 4\sqrt{2}$. Hence we have

Theorem 13: *Let*

$$r = lu + mv + nw \quad (|u| = 2, |v| = 2, |w| = 2)$$

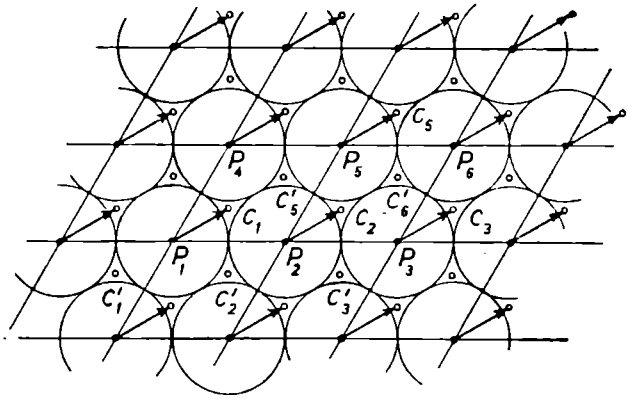


FIG. 19

be the lattice of the centres of a system $\{\sigma\}$ of unit spheres. For the density of packing of this kind of spherical system

$$\delta(\sigma) \leqq \delta(\sigma^{(1)}) = \frac{\pi}{6}\sqrt{2} \tag{3}$$

is true. The extremum is attained for an arrangement in the form of a lattice (2) *with the vectors*

$$u = (2, 0, 0), \quad v = (1, \sqrt{3}, 0), \quad w = (1, \tfrac{1}{3}\sqrt{3}, \tfrac{2}{3}\sqrt{6}). \tag{4}$$

There is an unproved conjecture that (3) also holds for arbitrary packings of unit spheres.

Each rhombus of the extremal lattice decomposes into two equilateral triangles with centroids C_ν and C_ν' (fig. 19).

In each parallelepiped, if one projects the edge parallel to the vector w onto the rhombus parallel to the (u, v)-plane, one obtains a line $P_\nu C_\nu$ which joins a vertex of the rhombus to the centroid. Displacing the P_ν by the vector $u = \overrightarrow{P_1 C_1}$ one obtains all these centroids C_ν. Therefore one can represent the extremal packing of spheres so that firstly one places a layer of spheres in the (u, v)-plane (with the angle $\alpha = 60°$) and then one imbeds the second layer

of spheres in the interspace denoted by the points C_ν. On the other hand, the centroids C'_ν of the other triangles have no spheres. If one applies the vector v again to the centres of the second layer one obtains the centres of the next layer, and so on.

However, one can also proceed in another way:—If one applies to the centres P_ν not the vector v but $-v$ then just the other points C'_ν (fig. 19) are reached by this displacement. One obtains a packing of equal density if one imbeds the spheres in the interspace denoted by C'_ν and it is possible in going from layer to layer to change from

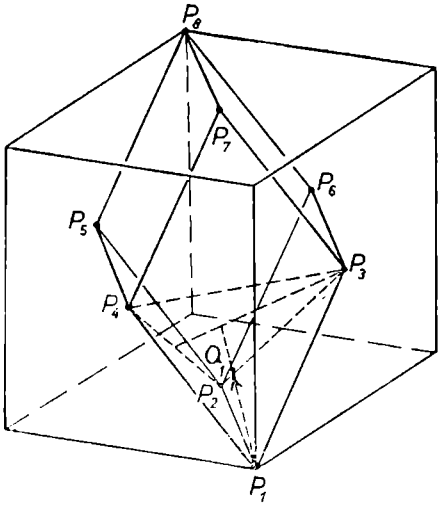


FIG. 20

v to $-v$ in a regular or also in quite a random way. In each case one obtains for the density of packing $d = \frac{\pi}{6}\sqrt{2}$. Of course, arrangements like these would no longer be referred to as being " lattice formed ".

For later applications we observe that one can also describe the arrangement of the centres in the extremal lattice as a " face-centred cubical lattice ". Moreover, we start from a cubical lattice

$$v = la + mb + nc. \tag{5}$$

The vectors a, b, c must be perpendicular to each other in pairs and have length $2\sqrt{2}$. If one adds to the lattice points (5) the centres of the side faces of each cube then one obtains an extremal lattice for the packing of unit spheres.

To see this, we single out a cube of the lattice (5) and, in addition, plot the centres of the faces (fig. 20). If one joins two opposite

vertices P_1 and P_8 of the cube with the centres P_2, P_3 and P_4 and P_5, P_6 and P_7 respectively of the adjoining side faces and also draws in the lines P_2P_6, P_3P_6, P_4P_5 and P_4P_7, then one obtains a parallelepiped which is inscribed in the cube and whose edges are all of length 2. Since P_2P_3 is also equal to 2, the whole face-centred cubical lattice is just the lattice of an extremal packing of spheres. From this fact one obtains once again the formula for the density of packing. Of course, one sees immediately that the unit spheres placed about the 6 face centres lie half inside the cube. One octant of each of the spheres belonging to the eight vertices of the cube lies inside the cube. There are $\frac{6}{2}+1 = 4$ full spheres. Therefore the density is

$$\delta = \frac{4 \cdot \frac{4}{3}\pi}{(2\sqrt{2})^3} = \frac{\pi}{6}\sqrt{2}.$$

2. The Finest Rigid Packing

By analogy with the packing of circles (p. 26), a spherical packing is said to be *rigid* if each sphere is touched by 4 others so that the 4 points of contact do not lie on a hemisphere. The extremal packing just investigated is certainly rigid: here, in fact, every sphere is touched by 12 others.

There exists a rigid packing of spheres in which each sphere is touched by exactly 4 others so that the points of contact form the vertices of a regular tetrahedron.

To describe this system we turn once again to the system $\{\mathfrak{s}^{(1)}\}$ with the densest packing. Suppose that a coordinate system is specified in the face-centred cube of fig. 20 so that zero coincides with P_1, the positive x-axis is in the direction from P_1 to P_2 and P_1P_4 lies in the xy-plane. Then the components of

$$\mathfrak{u} = \overrightarrow{P_1P_2}, \quad \mathfrak{v} = \overrightarrow{P_1P_4}, \quad \mathfrak{w} = \overrightarrow{P_1P_3}$$

are given in (4). Hence the vector $\mathfrak{x}$ from P_1 to the centroid Q_1 of the tetrahedron $P_1P_2P_3P_4$ has the components

$$\mathfrak{x} = \overrightarrow{P_1Q_1} = (1, \tfrac{1}{3}\sqrt{3}, \tfrac{1}{6}\sqrt{6}).$$

A new lattice of points $\{Q_\nu\}$ is defined by the displacement

$$\mathfrak{n} = \overrightarrow{P_1Q_\nu} = \overrightarrow{P_1P_\nu} + \mathfrak{x} = l\mathfrak{u} + m\mathfrak{v} + n\mathfrak{w} + \mathfrak{x} \quad (l, m, n \text{ integers}).$$

Then each point Q_ν is the centroid of a tetrahedron which is formed by 4 points of $\{P_\nu\}$: Q_1 of $P_1P_2P_3P_4$, Q_2 of $P_2P_3P_4P_5$ and so on. Conversely, each point of $\{P_\nu\}$ is the centroid of a regular tetrahedron whose vertices belong to the lattice $\{Q_\nu\}$. The union of the lattices $\{P_\nu\}$ and $\{Q_\nu\}$ is no longer a lattice in the sense of our definition, but rather a system of points which one can take as the centres of a rigid packing of spheres. One only needs to place a sphere of radius

$$\rho_1 = \tfrac{1}{2} \left| \varkappa \right| = \tfrac{1}{4}\sqrt{6}$$

about each point. We call this new system $\{\sigma^{(2)}\}$.

Exactly 4 of the spheres about the points of the system $\{Q_\nu\}$ lie wholly in the cube with sides $2\sqrt{2}$ (fig. 20), namely those with centres Q_1, Q_2, Q_3 and Q_4. The other spheres with centres Q_ν ($\nu>4$) lie completely outside the cube. To see this, one need only note that the vector $\varkappa = \overrightarrow{P_1Q_1}$ is in the direction of the principal diagonal of the cube. Since the spheres about the points P_ν and those about the points Q_ν have, of course, the same density then just 4 of the spheres about the points P_ν also lie in the cube. Therefore the density of our packing is

$$\delta_1 = \frac{8 \cdot \frac{4}{3}\pi\rho_1^3}{(2\sqrt{2})^3} = \frac{\pi\sqrt{3}}{16} \sim 0{\cdot}340.$$

Now this packing is by no means the finest of the rigid packings of spheres. Here, in the same way as for the plane (p. 26), one can find a new system in which one replaces each sphere of the system $\{\sigma^{(1)}\}$ of radius ρ_1 by 4 spheres with the following property:—These 4 spheres must have the same radius (ρ_2) and touch the spheres of $\{\sigma^{(1)}\}$ internally at their points of contact. In addition, ρ_2 must be chosen so that each of the 4 spheres is touched by the other three.

Since $2\alpha = 2\sin^{-1}\sqrt{\frac{2}{3}}$ is the angle between any two vectors from the centroid of the regular tetrahedron to its vertices, we have for the radius ρ_2 of the spheres of the new system $\{\sigma^{(2)}\}$

$$\sin\alpha = \sqrt{\tfrac{2}{3}} = \frac{\rho_2}{\rho_1 - \rho_2},$$

therefore

$$\rho_2 = \rho_1 \frac{1}{\sqrt{\frac{3}{2}}+1}.$$

Since each sphere of $\{\mathcal{S}^{(1)}\}$ is replaced by four smaller ones from $\{\mathcal{S}^{(2)}\}$, the following is true for the density δ_2 of the new system

$$\delta_2 = \frac{4 \cdot \frac{4}{3}\pi\rho_2^3}{\frac{4}{3}\pi\rho_1^3}\delta_1 = \frac{4}{(\sqrt{\frac{3}{2}}+1)^3}\delta_1 \sim 0{\cdot}123.$$

It is conjectured that this is the smallest value for the density of a rigid packing of spheres [II 3]. This has not yet been proved true.

Chapter V

LEBESGUE'S "TILE" PROBLEM

1. Formulation of the Problem

Many statments for the "tile problem" **(4)** (p. 2) remain valid when one considers any bounded point sets $\mathscr{S}$ instead of clusters of points (*finite* point sets). A segment of the plane (i.e. a closed domain of the plane) is called a *tile* if a point set of diameter 1 in the plane can be completely covered by it. By the diameter of any bounded point set we mean the least upper bound of the distance between any two points of the set. Our problem requires us to determine a tile of minimal surface area. The problem was first formulated in this form by Lebesgue in 1914 (see [V 3] p. 4). It is easy to show that a positive lower bound exists for the surface area of tiles (p. 65), but it is not immediately certain that there exists one (or more) tile(s) having area equal to this lower bound.

When investigating such existence problems it is convenient to restrict oneself to convex † tiles. Pál [V 3] specialised the Lebesgue problem in this way:—

(4a) *Let τ_0 be the lower bound of the surface area $\tau(\mathscr{T})$ of all convex tiles.*

(a) *The value of τ_0 is to be determined.*

(b) *It is to be ascertained whether there exists a tile of $\mathscr{T}$ with area equal to τ_0.*

(c) *If there exist tiles with minimal area, are all tiles $\mathscr{T}$ with surface measure τ_0 of the form to be determined later?*

Since we have defined the tiles to be closed sets, we can also assume the sets $\mathscr{S}$ to be closed. In Sections 1-5 of this chapter we shall always indicate by the letter $\mathscr{S}$ a closed set of diameter 1.

† A domain is said to be convex if, when the points P and Q belong to the domain, the line PQ also belongs to it.

We want to approach our problem first of all by compiling examples of tiles and trying to find those with the smallest possible area. It is easy to show that the square with side of length 1 is a suitable tile for each set $\mathscr{S}$. To prove this we introduce the concept of the *support line*. This is a line l which contains at least one point of $\mathscr{S}$ and is such that a half plane to one side of it contains no point of $\mathscr{S}$. Then, to each line l in the plane of $\mathscr{S}$, there exist two parallel support lines l_1 and l_2. They are at a distance apart d_1 which is at most equal to 1. Then, if l_3 and l_4 are the support lines perpendicular to l_1 and at a distance apart $d_2 \leqq 1$, $\mathscr{S}$ is enclosed by these 4 lines, forming a rectangle with sides of length $d_1 \leqq 1$ and $d_2 \leqq 1$. Thus it is clear that the square with side of length 1 is a tile. We denote this tile by $\mathscr{T}_2$. We want to keep the suffix 1 for the Jung circle of radius $r = \frac{1}{3}\sqrt{3}$. Its surface area $A = \pi/3$ is somewhat greater than 1 ($A = 1{\cdot}047197\ldots$).

It is immediately obvious that no circle with radius $r < \frac{1}{3}\sqrt{3}$ is a tile:—The vertices of an equilateral triangle with side of length 1 forms a cluster of points of diameter 1, the closed triangular region forms a convex set $\mathscr{S}$. It is covered by its circumscribed circle ($r = \frac{1}{3}\sqrt{3}$) but not by any smaller circle. One can easily prove directly that $\mathscr{T}_1$ actually is a tile (see e.g. [V 4]). We shall obtain this property for $\mathscr{T}_1$ from the fact that the regular hexagon inscribed in Jung's circle is a tile.

2. The Hexagonal Tile

To obtain another tile $\mathscr{T}_3$ we enclose $\mathscr{S}$ in an equiangular hexagon with sides s_ν ($\nu = 1, 2, \ldots, 6$) according to the following rule:—The sides s_1 and s_4 must lie on parallel lines of support whose distance apart is 1. s_2 and s_3 are on lines of support which enclose with s_1 or s_4, respectively, an angle of 120°. Finally s_5 and s_6 must lie on parallels to s_2 or s_3, respectively, at a distance of 1 unit. Therefore s_5 and s_6 *can* be lines of support but *need* not be (fig. 21). Numbering the sides must produce a positive contour direction.

By projecting the sides s_1 and s_2 onto a line which intersects s_3 and s_6 at right angles, one obtains

$$(s_1 + s_2) \cos 30° = 1.$$

In the same way

$$s_2 + s_3 = s_3 + s_4 = s_4 + s_5 = s_5 + s_6 = s_6 + s_1 = \frac{2}{\sqrt{3}},$$

and hence it follows that

$$s_2 = s_4 = s_6(= s), \quad s_1 = s_3 = s_5(= \sigma).$$

If $s_1 = s_2$, then $s = \sigma$ and the equiangular hexagon is regular. We want to show that in every case a regular hexagon (with distance 1 between opposite sides) can be used as a covering of $\mathscr{S}$.

Let $s_1 > s_2$, say. Then we introduce a notation for direction in the plane of $\mathscr{S}$. We associate the direction $\phi = 0$ with the base line s_1 and the direction $\phi = 60°$ with the line s_2. For each direction ϕ with $0 \leqq \phi \leqq 60°$, there can now be found a parallel strip which is constructed from a line of support and a parallel at a distance of one

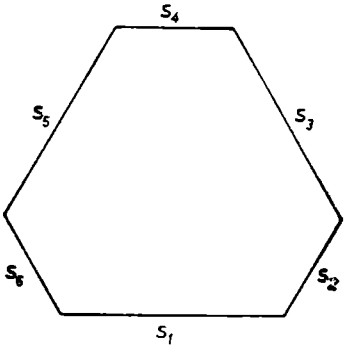


FIG. 21

unit and which encloses $\mathscr{S}$. The process just described motivates constructing an equiangular hexagon enclosing $\mathscr{S}$. Let the sides of the hexagon be $s_n(\phi)$, $n = 1, 2, \ldots, 6$. Then

$$S(\phi) = s_1(\phi) - s_2(\phi)$$

is a continuous function of ϕ. For $\phi = 0$ we have by hypothesis $S(0) = s_1 - s_2 > 0$. On the other hand

$$S(60°) = s_1(60°) - s_2(60°) = s_2 - s_3 = s_2 - s_1 < 0.$$

Therefore, between $\phi = 0$ and $\phi = 60°$, by continuity, there must be a value of ϕ for which $S(\phi)$ is zero. Then $s_1(\phi) = s_2(\phi)$ and the corresponding equiangular hexagon is regular.

Hence it has been shown that the regular hexagon with in-circle of radius $\rho = \frac{1}{2}$, therefore with circumcircle of radius $r = \frac{1}{3}\sqrt{3}$, is a tile. We denote it by $\mathscr{T}_3$. Of course the circumcircle of $\mathscr{T}_3$ is also a tile. This is just Jung's circle which we have denoted by $\mathscr{T}_1$.

In this way we have found 3 tiles whose areas decrease with the suffix n:—

$$\tau_1 = F(\mathscr{T}_1) = 1{\cdot}04797\ldots > \tau_2 = F(\mathscr{T}_2) = 1 > \tau_3 = F(\mathscr{T}_3) = \tfrac{1}{2}\sqrt{3} = 0{\cdot}8660.$$

3. The Tiles of Pál and Sprague

The tile $\mathscr{T}_3$ can easily be improved on. If one draws to the incircle of $\mathscr{T}_3$ the tangents which are perpendicular to the diagonals, one obtains at the 6 vertices V of $\mathscr{T}_3$ (fig. 22) 6 isosceles triangles

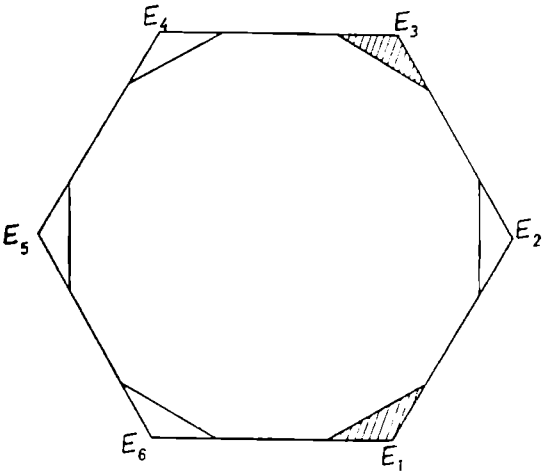


FIG. 22

Δ_ν. Now let a point set $\mathscr{S}$ be covered by $\mathscr{T}_3$. Then all the triangles Δ_ν cannot be occupied † by points of $\mathscr{S}$:—The interior points of opposite triangles are at a distance apart which is greater than 1. Therefore if Δ_1 is occupied then Δ_4 is free and conversely. Similarly, both Δ_2 and Δ_5 and Δ_3 and Δ_6 respectively cannot be occupied simultaneously. Therefore there are the following 8 possibilities for 3 corners to be occupied:—

$$\begin{array}{lll} \Delta_1, \Delta_2, \Delta_3 & \Delta_4, \Delta_5, \Delta_6 & \\ \Delta_2, \Delta_3, \Delta_4 & \Delta_5, \Delta_6, \Delta_1 & \Delta_1, \Delta_3, \Delta_5 \\ \Delta_3, \Delta_4, \Delta_5 & \Delta_6, \Delta_1, \Delta_2 & \Delta_2, \Delta_4, \Delta_6. \end{array} \qquad (1)$$

The other three respectively are free. It is immediately obvious that, of the free corners, there are always two whose suffixes differ by

† We say that a triangle Δ_ν is " occupied " if a point of $\mathscr{S}$ lies in the interior of the triangle or on a point of the boundary of Δ_ν which does not belong to the tangent.

two or four. Hence $\mathscr{T}_3$ does not cease to be a tile if one cuts off two corners which are not adjacent and not opposite, say Δ_1 and Δ_3. The remaining tile $\mathscr{T}_4 (= ABV_2CDV_4V_5V_6$ in fig. 23) given by PÁL [V 3] has area

$$\tau_4 = F(\mathscr{T}_4) = 2 - \tfrac{2}{3}\sqrt{3} = 0{\cdot}845299\ldots.$$

R. SPRAGUE [V 5] has shown that one can also improve on this tile. One can show this as follows:—Let $\mathscr{S}$ be a point set of diameter

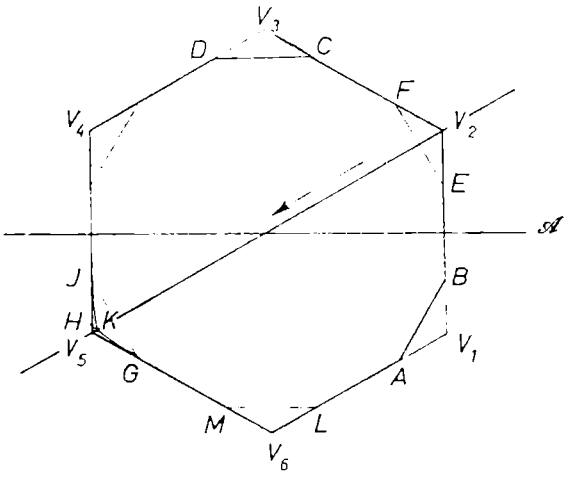


FIG. 23

1 which is covered by $\mathscr{T}_4$. Only one of the two isosceles triangles at V_2 and V_5, $\Delta_2 = V_2EF$ and $\Delta_5 = JGV_5$, can be occupied.† First of all we suppose that a point $P \in \mathscr{S}$ lies on the line CF. Then we draw about C the circular arc GH with radius 1 which must touch V_5V_6 in G and intersect V_4V_5 in H. Then no point of $\mathscr{S}$ can lie in the part of the tile cut off by this arc since all these points are at a distance greater than 1 from the point of $\mathscr{S}$ lying on CF. We want to show that we achieve our purpose in each case with the residual tile left when $\Delta = V_5GH$ (GH being a circular arc) is cut off $\mathscr{T}_4$.

Therefore, let us suppose that no point of $\mathscr{S}$ lies on CF. If one does lie on EF then the whole triangle Δ_5 is unoccupied and we certainly achieve our purpose with the reduced tile $\mathscr{T}_5 = \mathscr{T}_4 - \Delta$. If a point of $\mathscr{S}$ lies on EB or AB, one can rotate the tile about the axis V_2V_5 (fig. 23) and hence we have that CF or DC respectively lies on a point of $\mathscr{S}$. In the first case—we have already seen this—we certainly achieve our purpose with the "residual tile". In the

† F must lie on V_2V_3, G on V_5V_6.

second case we displace the tile in the direction of the axis $\mathscr{A}$ to the left until a point of $\mathscr{S}$ lies on the polygon *CFEBAL*. If the broken line *CFEB* is occupied then the residual tile forms the covering. If, however, the displacement leads to *BAL* being occupied then we rotate the tile about the axis $\mathscr{A}$ and so have *EFC* occupied. This rotation is permitted if *CD* is occupied, since, in this case, the opposite corner ($\Delta_6 = MLV_6$) must be free.

Now let us suppose that in a covering of a set $\mathscr{S}$ by $\mathscr{T}_4$ the whole broken line *ABEFCD* (and, of course, the triangle Δ_2 also) is free of points of $\mathscr{S}$. Then one only needs to displace the whole tile in the direction V_2V_5 (arrow in fig. 23) until a point of *ABEFCD* reaches a point of $\mathscr{S}$. Carrying out this displacement does not cancel the covering of the other points of $\mathscr{S}$ by the tile. Hence this general case has also led back to that already discussed and we have shown that $\mathscr{T}_5 = \mathscr{T}_4 - \Delta$ is also a tile.

$\mathscr{T}_5$ is not symmetric about the axis V_2V_5. This suggests that one can also cut off another piece of the triangle Δ_5 by the circle with centre B and radius $BJ = 1$ (fig. 23). Let K be the point of intersection, lying on V_2V_5, of the circles about B and C with radius 1. Then, indeed, $\mathscr{T}_6 = ABV_2CDV_4JKGV_6$ is a tile, as we proceed to show in the next section.

4. Curves and Domains of Constant Width

A bounded domain is said to be of bounded width b if each pair of parallel lines of support has the same distance apart b. The boundary of such a domain is called a curve of constant width. The circular disk of diameter b is, of course, a domain of this kind. Other simple examples are certain triangles with circular arcs as sides. If one draws about each vertex of a $(2n+1)$-gon the circular arcs passing through the vertices of the opposite side one obtains a domain of constant width d, where d is equal to the radius of the circular arcs (fig. 24 for $2n+1 = 5$).

Now the following theorem is significant for our tile problem:—

Theorem 14: *Each cluster of points $\mathscr{C}$ is a subset of a closed, convex domain with constant width* 1.

To prove this we place about $\mathscr{C}$ a polygon $\mathscr{P}$ which is constructed from support lines of this set and whose vertices are points of the cluster $\mathscr{C}$. Let $\mathscr{C} = \{Q_1, Q_2, \ldots, Q_m\}$ be the cluster of points and $P_1P_2P_3\ldots P_n$ the polygon, where $\{P_1, P_2, \ldots, P_n\}$ is a subset of $\mathscr{C}$.

The distance between at least one pair of points $P_\nu P_\mu$ must then be equal to 1. Let this be the case for P_1P_i.

We denote by $\mathcal{H}_1$ the half-plane determined by the line P_1P_i and containing the points $P_2, P_3, \ldots, P_{i-1}$; let the other half-plane be $\mathcal{H}_2$ (fig. 25).

Let us now draw circles $(c_1, c_2, \ldots, c_i)$ with centres $P_1, P_2, P_3, \ldots, P_i$ and radius 1. c_j and c_{j+1} intersect at the point R_j of the half-plane $\mathcal{H}_2$. In this way there results in $\mathcal{H}_2$ a polygon $P_iR_1R_2\ldots R_{i-1}P_1$ with circular arcs as sides (curvilinear polygon). Now let us also draw a circle with centre R_j and radius 1 then we also have in the half-plane $\mathcal{H}_1$ a curvilinear polygon $P_1P_2P_3\ldots P_i$ (fig. 25).

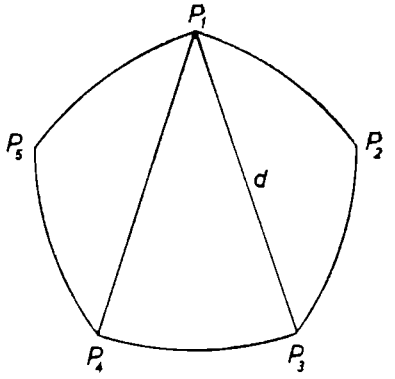


FIG. 24

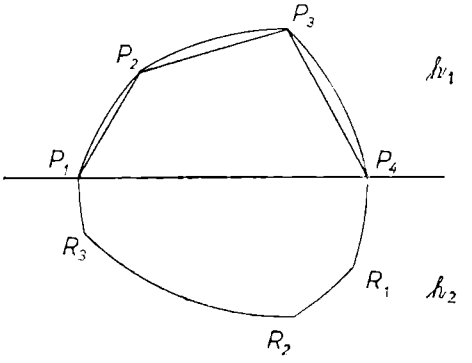


FIG. 25

Then the whole curvilinear polygon $P_1P_2\ldots P_iR_1\ldots R_{i-1}P_1$ is obviously a curve of constant width 1:—One of a pair of parallel support lines 1 unit apart always goes through a vertex of the curvilinear polygon and the other is a tangent to the opposite arc. All the points of our cluster $\mathscr{C}$ must lie in the domain determined by the curvilinear polygon since those lying in $\mathcal{H}_2$ must also be at a distance of at most 1 unit from all points P_j of the half-plane $\mathcal{H}_1$.

The convex domain of width 1 constructed in this way is also called the Reuleaux curvilinear polygon belonging to $\mathscr{C}$.

Theorem 14 can now be used to prove that $\mathscr{T}_6$ is a tile. Let $\mathscr{C}$ be a cluster of points of diameter 1 and $\mathscr{R}$ a Reuleaux curvilinear polygon associated with it. Then $\mathscr{R}$ can be covered by the tile $\mathscr{T}_4$. Since there is a pair of parallel lines of support 1 unit apart corresponding to each direction, a boundary point of $\mathscr{R}$ must lie on each of the lines BV_2, V_2C, DV_4, V_4V_5, V_5V_6 and V_6A (fig. 23). One at least, of the triangles Δ_2 and Δ_5 must remain free of points. If Δ_5

is free then the smaller tile $\mathscr{T}_6$ is certainly sufficient to cover $\mathscr{C}$. Therefore let us suppose that $\Delta_2 = EV_2F$, rather than Δ_5, is free of points of $\mathscr{R}$. The point $\mathscr{R}$ lying on V_2C must therefore lie on FC and that lying between B and V_2 must lie on BE. Then no further points of $\mathscr{R}$ can lie outside of the circle with centre C and radius $CG = 1$ and with centre B, radius $BJ = 1$. However, this means that the tile $\mathscr{T}_6$ is also sufficient to cover $\mathscr{R}$ and hence also $\mathscr{C}$.

In this way it has been shown, in the first instance, only that every *cluster* of points $\mathscr{C}$ can be covered by $\mathscr{T}_6$, but not unconditionally every arbitrary set of diameter 1. Now one can easily extend the statement of Theorem 14 to any set of diameter 1. But it is even simpler to deduce directly the tile property for an arbitrary set $\mathscr{S}$ from that for a cluster of points.

For this purpose let us think of the points, enumerated in any way, of a subset which is everywhere dense in $\mathscr{S}$: $P_1, P_2, \ldots, P_n, \ldots$. Then let $\mathscr{C}_n$ be the cluster of points $\{P_1, P_2, \ldots, P_n\}$. Each cluster $\mathscr{C}_n$ is covered by $\mathscr{T}_6$. Let $\mathscr{T}_6^{(n)}$ be the domains which are congruent to our tile $\mathscr{T}_6$ thought of as movable and which cover the corresponding clusters $\mathscr{C}_n$ with the same suffix. From the sequence $\mathscr{T}_6^{(n)}$ one can now select a sub-sequence $\mathscr{T}_6^{(n)'}$ for which the (finitely many) vertices $\mathscr{T}_6^{(n)'}$ converge to limit points. These limit points determine a new position $\mathscr{T}_6^*$ of our tile $\mathscr{T}_6$. It obviously covers all the points P_n and hence also $\mathscr{S}$ since each point of $\mathscr{S}$ is a point of accumulation of points of the set $\{P_n\}$.

The tile $\mathscr{T}_6$ (also discovered by R. SPRAGUE) has surface area

$$\tau_6 = A(\mathscr{T}_6) = 0{\cdot}844144\ldots.$$

In order to estimate the convergence of this tile,† the best to our knowledge at present, to the lower limit τ_0, we now try to find lower limits for the area of all tiles.

Each tile which is to cover all sets of diameter 1 must contain a circle of this diameter. The area of each tile is therefore at least equal to $\pi/4$. In order to improve this limit we consider that each tile must also cover an equilateral triangle with sides of length 1.

† One sees immediately that one cannot cut off any of the vertices V_2, V_4 and V_6 of the tile $\mathscr{T}_6$ (Fig. 23) because they are required to cover the equilateral triangle with sides of length 1. R. SPRAGUE conjectures that the absolute necessity of the other vertices of the tile can also be proved by suitable, simple clusters of points.

However, even if one cannot cut off any more from $\mathscr{T}_6$, it is still possible that there could exist yet another kind of tile with smaller area.

Therefore each convex tile contains the convex hull of a union of sets of circle and triangle. PÁL has shown [V 3] that this set has minimal surface area if the centre of the circle lies at the centroid of the triangle. Therefore, the surface area of this figure $\mathscr{F}$ is a lower limit for the surface area of all tiles. *However, $\mathscr{F}$ itself is not a tile.* In fact one can show that $\mathscr{F}$ does *not* cover the regular pentagon with constant width 1 (fig. 24). Hence $A(\mathscr{F}) = 0{\cdot}825711\ldots$ is † smaller than the lower bound τ_0 of the areas of all the tiles and we have as the prevailing result

$$0{\cdot}825711\ldots < \tau_0 \leqq \tau_6 = A(\mathscr{T}_6) = 0{\cdot}844144\ldots$$

or

$$\tau_0 = 0{\cdot}834928 \pm 0{\cdot}009217\ldots. \tag{2}$$

Hence the numerical value mentioned in problem (**4a**) (p. 57) is in fact not determined but is at least bounded. The question (b) in problem (**4a**) is to be answered in the affirmative. One can show using Blaschke's theorem of choice ‡ that there actually exists at least one minimal convex tile. We will not go into the proof of this existence theorem since it is rather involved though not difficult.§ The problem of the practical determination of one or more minimal tiles is still open.

Theorem 15: *Among all the convex tiles which cover each plane domain of diameter* 1 *there exists at least one tile with the minimal surface area τ_0. For this numerical value*

$$0{\cdot}825711\ldots < \tau_0 \leqq \tau_6$$

is true, where τ_6 is the area of Sprague's tile $\mathscr{T}_6$.

5. Applications

In 1933 G. BORSUK ‖ posed the question whether it is always possible to decompose a point set of n-dimensional space of diameter 1 into $n+1$ parts so that the diameter of each subset is less than 1.

One can immediately give an affirmative answer to this question for the plane using our results on the tile problem. Each plane set of diameter 1 can be completely covered by the hexagonal tile $\mathscr{T}_3$. If one can divide this tile into 3 parts, each with diameter smaller

† Here the positive answer to question (b) of problem (**4a**) (p. 57) has been anticipated.

‡ See e.g. [V 10].

§ See e.g. [V 3].

‖ *Fund. Math.*, **20**, pp. 177-190, 1933.

than 1, then a corresponding decomposition is possible for each set of diameter 1.

To divide the tile $\mathscr{T}_3$ in the required way, from the centre of this tile we drop the perpendiculars onto three, non-adjacent sides. Let the feet of these perpendiculars which enclose angles of 120° be P_1, P_2 and P_3. $\mathscr{T}_3$ is divided by these perpendiculars into 3 sub-tiles of diameter $P_\nu P_\mu = \frac{1}{2}\sqrt{3}$. Hence we have an affirmative answer to Borsuk's problem for $\mathscr{R}_2$.

Theorem 16: *Each plane set of diameter* 1 *can be decomposed into* 3 *subsets whose diameters are at most equal to* $\frac{1}{2}\sqrt{3}$.

Another application of our results for the covering problem leads to the

Theorem 17: *Let* $\mathscr{D}$ *be a set of open disks of radius* 1, *each two of which have a point in common. Then there is an equilateral triangle PQR with sides of length* 1 *such that each disk contains at least one of the points P, Q or R.*

We can visualise this theorem as follows:—One imagines the disks of the set $\mathscr{D}$ cut out of paper and laid on a drawing board. Then one can fix the disks by pins in 3 points P, Q and R (which form the vertices of an equilateral triangle with sides of length 1) so that each disk is fastened to the board by (at least) one pin.

To prove this theorem we note that the centres of the disks of $\mathscr{D}$ form a cluster of points $\mathscr{C}$ with diameter $d<2$, since each two disks have a point in common. $\mathscr{C}$ can be covered by a tile of type $\mathscr{T}_3$ for which the distance between opposite sides of the hexagon is equal to 2. Let us denote the vertices of this tile by A_ν ($\nu = 1, 2, \ldots, 6$) and the midpoints of A_1A_3, A_3A_5 and A_5A_1 by P, Q and R. Then PQR is an equilateral triangle with sides of length 1. Since each point of the tile is at a distance of at most 1 unit from at least one of the 3 points P, Q and R, this is also true for the centres of the disks which are covered by the tile. This means that every disk of $\mathscr{D}$ covers at least one of the three points.

One can generalise the problem considered here and ask by how many points can a set $\mathscr{D}'$ of disks be fixed if each two disks of this set have an (interior) point in common but the radii may be *different*. One knows that 5 points suffice † and L. FEJES TOTH conjectures that one can also manage with 4 points.

† See HADWIGER in Ungelöste Probleme (Unsolved Problems), No. 19, *El. d. Math.*, XII, pp. 109-110, 1957.

6. The Covering Problem in $\mathscr{R}_3$

Jung showed in 1903 [V 1] that every cluster of points of diameter 1 in three dimensional space can be placed in a sphere of radius $r = \frac{1}{2}\sqrt{\frac{3}{2}}$. Kirsch gave a new, elementary proof [V 7]. This result suggests the problem of finding *a covering body (D-body) of smallest volume for a cluster of points of diameter* 1 *of* $\mathscr{R}_3$. Jung's sphere is a D-body $\mathscr{D}_1$ with volume $V(\mathscr{D}_1) = \frac{\pi}{4}\sqrt{\frac{3}{2}} \sim 0{\cdot}95\ldots$. We now prove

Theorem 18: *The regular octahedron with edges of length* $a = \sqrt{\frac{3}{2}}$ *is a D-body* $\mathscr{D}_2$ *with volume*

$$V(\mathscr{D}_2) = \tfrac{1}{2}\sqrt{3} \sim 0{\cdot}866\ldots.$$

To prove † this theorem we need the concept of the "plane of support" of a cluster of points $\mathscr{C}$. By this we mean (in analogy to the line of support, see p. 58) a plane η which has at least one point in common with $\mathscr{C}$ and such that one of the two half-spaces determined by η does not contain any points of $\mathscr{C}$.

Now let $A_1A_2A_3A_4B_1B_2$ be an octahedron $\mathcal{O}$ whose axes A_1A_3, A_2A_4 and B_1B_2 coincide with the x-, y- and z-axes of a coordinate system. Let the edges be chosen of length 1, say. We use this octahedron simply to have a convenient notation for certain planes and angles. First of all we denote the planes determined by the points $A_\nu A_{\nu+1}B_1$ by ε_ν ($\nu = 1, \ldots, 4, A_5 = A_1$). Then the three planes ε_1, ε_3 and ε_4 form at B_1 a solid angle $\mathscr{A}$ containing the octahedron.

Then there are 3 well-defined planes of support η_1, η_3 and η_4 of $\mathscr{C}$ which are parallel to ε_1, ε_3 and ε_4 and at a point S form a solid angle $\mathscr{A}_1$ which is parallel to the angle $\mathscr{A}$ and contains $\mathscr{C}$. Now let us suppose that the plane η_2 through S and parallel to ε_2 is not a plane of support of $\mathscr{C}$ (fig. 26). Then there are two possibilities:—

1. All the points $C \in \mathscr{C}$ lie inside the 4-sided angle $\mathscr{A}_2$ which is formed at S by the planes η_μ ($\mu = 1, 2, 3, 4$). Then at least one point C lies on the sides of $\mathscr{A}_2$ determined by the planes η_1, η_3 and η_4 but none on η_2.
2. There is at least one point $C \in \mathscr{C}$ which lies outside $\mathscr{A}_2$.

First of all let us consider the first case. There exists for this position of $\mathscr{A}_2$ a plane of support η_2' of $\mathscr{C}$ which is parallel to η_2.

† This theorem was stated without proof by D. Gale [V 6].

It must go through a point S' of the edge of $\mathscr{A}_2$ determined by η_1 η_4 (fig. 26).

We now have at S' a new solid angle $\mathscr{A}_3$ which is located parallel to the angle $\mathscr{A}_2$. It is formed by the planes η_1, η_2', η_4 and a plane η_3' parallel to η_3. Now, of these 4 planes η_1, η_2' and η_4 are planes of support while η_3' does not have this property.

We now take an interior point of $\mathscr{A}_3$ as the origin of a coordinate system. Then we can introduce the *directed* distance of a point P from a plane using the Hesse standard forms of the plane. With this choice of origin the distance from the planes η_1, η_2', η_3', η_4 for

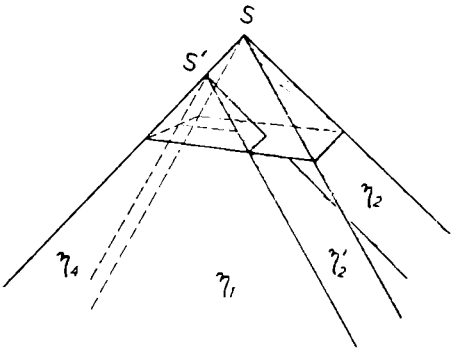


FIG. 26

all interior points of $\mathscr{A}_3$ becomes negative. Now let d be the maximum of the directed distance of the points $C_\nu \in \mathscr{C}$ from η_2 and d' that from η_3'. Then in our case $d<0$ but $d'>0$ since at least one point $C_\nu \in \mathscr{C}$ lies on the plane η_3 placed outside the angle $\mathscr{A}_3$.

Now let us imagine the octahedron $\mathcal{O}$ replaced by an octahedron $\mathcal{O}(\alpha) = A_1^\alpha A_2^\alpha A_3^\alpha A_4^\alpha B_1 B_2$ which results from $\mathcal{O} = \mathcal{O}(0)$ by rotation through the angle α about the axis B_1B_2. Let the corresponding planes be $\varepsilon_\mu(\alpha)$. Here a distance $d(\alpha)$ depending on α corresponds to the maximal distance d from η_2. $d(\alpha)$ is certainly a continuous function of α. Now, for $\alpha = \pi/2$, we get a rotation of the octahedron through a right angle so that

$$\varepsilon_\mu\left(\frac{\pi}{2}\right) = \varepsilon_{\mu+1}, \quad \varepsilon_5 = \varepsilon_1.$$

Then the solid angle $\mathscr{A}_2(\pi/2)$ corresponding to this octahedron has the property that $\eta_1(\pi/2)$, $\eta_3(\pi/2)$ and $\eta_4(\pi/2)$, and therefore η_2', η_4 and η_1, are the planes of support of $\mathscr{C}$. However, the angle at S' denoted just above by $\mathscr{A}_3$ has this very property. Therefore we

have $\mathscr{A}_2(\pi/2) = \mathscr{A}_3$ and correspondingly $d(\pi/2) = d'$. Hence

$$d(0) = d < 0, \quad d\left(\frac{\pi}{2}\right) = d' > 0.$$

From continuity considerations there must therefore be an angle α for which $d(\alpha) = 0$. Then we have a solid angle $\mathscr{A}_2(\alpha)$ which contains $\mathscr{C}$ completely and has at least one point of $\mathscr{C}$ in each side.†

If there is a point $C_\nu \in \mathscr{C}$ lying outside $\mathscr{A}_2$, one comes to the same result in the end. In the second case one starts with a parallel displacement of $\mathscr{A}_2$ into a position $\mathscr{A}'_3$ for which three sides belong to planes of support while the maximal distance d'' from the last side is negative. The details of this argument are left to the reader to supply. In this case also one reaches the conclusion that there exists a solid angle $\mathscr{A}'_2(\alpha)$ for which $d'(\alpha)$ is zero.

Therefore in each case a 4-sided solid angle exists which is congruent to the 4-sided angle of the octahedron at B_1 and contains on each of its sides at least one point of $\mathscr{C}$. However $\mathscr{C}$ has diameter 1. In each direction there are therefore two parallel planes at distance 1 unit apart, of which at least one is a plane of support and between which the whole cluster of points $\mathscr{C}$ must lie. Then the angle $\mathscr{A}_2(\alpha)$ (or $\mathscr{A}'_2(\alpha)$ respectively) and the planes parallel to its sides and 1 unit apart determine an octahedron which contains $\mathscr{C}$. From the fact that the opposite sides are 1 unit apart, one determines immediately the length of the edges $a = \sqrt{\frac{3}{2}}$. Thus Theorem 18 has been proved.

If one places on each of the 8 surfaces of the octahedron a tetrahedron with edges of length $a = \sqrt{\frac{3}{2}}$ then one obtains a tetrahedron with edge of length $a_1 = 2\sqrt{\frac{3}{2}}$ which is also of course a D-body. In this way a theorem of D. GALE [V 6] for $\mathscr{R}_3$ is proved: D. GALE has shown that in n-dimensional space each cluster of points of diameter 1 can be enclosed in a simplex with edge of length $\left(\frac{n(n+1)}{2}\right)^{\frac{1}{2}}$. However, a simplex is not a particularly good D-body since, from the results of GALE, the edge length and hence also the volume increases with n without bound. However, for Jung's sphere this is not the case: the generalisation of Jung's theorem to

† This condition is also satisfied for two sides of the angle if at least one point lies on the common edge.

n-dimensional space gives for the radius (see e.g. [VI 11] p. 140)

$$r_n = \sqrt{\frac{n}{2n+2}}, \quad \lim_{n\to\infty} r_n = \tfrac{1}{2}\sqrt{2}.$$

It is possible to reduce the D-body $\mathscr{D}_2$ by cutting off parts in the same way as for the hexagonal tile $\mathscr{T}_3$ in the plane. To see this we place at a distance $\frac{1}{2}$ planes parallel to the planes of symmetry which meet $\mathscr{D}_2$ at 4 vertices and these cut off from $\mathscr{D}_2$ at each vertex a half octahedron with height $h = \frac{1}{2}(\sqrt{3}-1)$. We denote these 4-sided pyramids by $\mathscr{P}_\nu$ ($\nu = 1, \ldots, 6$). They may belong to A_1, A_2, A_3, A_4, B_1, B_2 respectively. If, in the covering of a cluster of points $\mathscr{C}$ by $\mathscr{D}_2$, a point of $\mathscr{C}$ lies † in one of the pyramids $\mathscr{P}_\nu$ then the opposite pyramid must be unoccupied since the distance between the bases of opposite pyramids is exactly equal to 1. Therefore of the six pyramids at most three can be occupied.

If one notes all the possible combinations of free and occupied pyramids then one knows at once that in each case there remain unoccupied three pyramids which belong to the vertices of a triangular face of $\mathscr{D}_2$. However, this means that there is a D-body $\mathscr{D}_3$ which arises from $\mathscr{D}_2$ by cutting off the pyramids $\mathscr{P}_1$, $\mathscr{P}_2$ and $\mathscr{P}_5$. Since the volume of each pyramid cut off is equal to $\frac{1}{6}(3\sqrt{5}-5)$, we have for the volume of $\mathscr{D}_3$

$$V(\mathscr{D}_3) = V(\mathscr{D}_2) - 3 \cdot \tfrac{1}{6}(3\sqrt{3}-5) \sim 0{\cdot}768\ldots. \qquad (3)$$

Theorem 19: *If one cuts off from the D-body $\mathscr{D}_2$ three right pyramids at vertices (which are not opposite each other) and of volume $\frac{1}{6}(3\sqrt{3}-5)$ then one again obtains a D-body $\mathscr{D}_3$ with volume* (3).

It is certainly possible to reduce this D-body still further by carrying over to $\mathscr{R}_3$ the arguments of R. SPRAGUE. The problem of the D-body of smallest volume in $\mathscr{R}_3$ is unanswered just as that concerning the smallest tile in the plane.

One can use the D-body $\mathscr{D}_2$ to prove the Borsuk conjecture (p. 65) for $\mathscr{R}_3$. B. GRÜNBAUM [V 11] has ‡ shown recently, using the covering property of $\mathscr{D}_2$, that one can decompose each point set of diameter 1 in $\mathscr{R}_3$ into 4 subsets each of which has diameter $d \leqq 0{\cdot}88$. The question whether one can improve this result by using a D-body of smaller volume remains open.

† Here the points of the bases of the pyramids $\mathscr{P}_\nu$ are not considered as belonging to $\mathscr{P}_\nu$.

‡ See in this connection and for further basic problems of the theory of content [II 1], [VI 2], [VI 11] and [VI 12].

Chapter VI

ON THE EQUIVALENCE OF POLYHEDRA BY DISSECTION

1. The Plane Problem

In order to lead up to the problems of the theory of content in $\mathfrak{R}_3$ (problem (5) p. 2), it is expedient to treat first of all the corresponding questions in the plane.

One can associate with the triangle Δ the functional $F(\Delta) = \frac{1}{2}gh$ (g: base, h: height) as "content". Thus one can interpret the product $\frac{1}{2}gh$ either as a product of measures or in the sense of Hilbert's line multiplication as a product of lines.† One can easily show that this product is independent of the special choice of base. For a simple, closed, plane polygon $\mathscr{P}$ one can assign as area $F(\mathscr{P})$, the sum of the areas of the constituent triangles which result from any triangulation of the polygon. Hence $F(\mathscr{P})$ is independent of the particular kind of triangulation.

The functional $F(\mathscr{P})$ defined in this way has the following properties:—

1. $F(\mathscr{P}) = F(\mathscr{Q})$, $\mathscr{P} \equiv \mathscr{Q}$.
2. $F(\mathscr{P}+\mathscr{Q}) = F(\mathscr{P})+F(\mathscr{Q})$ if $\mathscr{P}$ and $\mathscr{Q}$ have no interior points in common.
3. $F(\mathscr{U}) = 1$ for the unit square $\mathscr{U}$.
4. From $F(\mathscr{P}) = F(\mathscr{Q})$ it follows that $\mathscr{P} \overset{d}{=} \mathscr{Q}$ (see footnote ‡).

The last property is particularly noteworthy because, as a result of it, the "content" has a very clear geometrical meaning:— Polygons of equal content can always be decomposed into *finitely many triangles which are congruent in pairs.*

† In connection with this and for further basic problems of the theory of content see [II 1], [VI 2] and [VI 12].

‡ The equivalence by dissection was defined already on p. 3.

To prove this important property 4. of our functional, we note the following rules of equivalence by dissection (see e.g. [VI 2], p. 134, *et seq.*):—

(a) From $\mathscr{P} \overset{d}{=} \mathscr{Q}$ and $\mathscr{Q} \overset{d}{=} \mathscr{R}$ it follows that $\mathscr{P} \overset{d}{=} \mathscr{R}$,

(b) From $\mathscr{P} = \mathscr{P}_1 + \mathscr{P}_2$, $\mathscr{Q} = \mathscr{Q}_1 + \mathscr{Q}_2$, $\mathscr{P}_1 \overset{d}{=} \mathscr{Q}_1$, $\mathscr{P}_2 \overset{d}{=} \mathscr{Q}_2$ it follows that $\mathscr{P} \overset{d}{=} \mathscr{Q}$.

We prove first of all the following lemma:—

Parallelograms with the same base and height are equivalent by dissection.

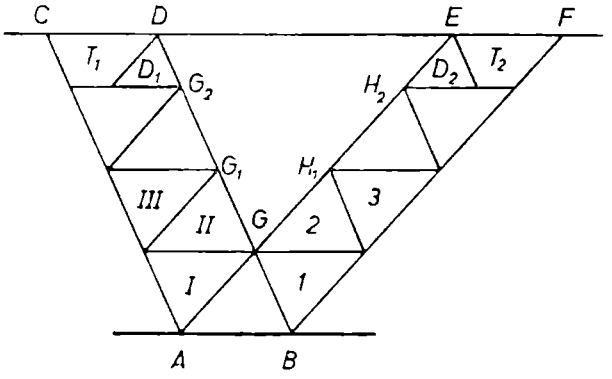


FIG. 27

The proof of this lemma is particularly simple if, in the parallelograms $ABCD$ and $ABEF$, the point E lies between C and D. Then obviously

$$AEDB \equiv AEDB, \quad AEC \equiv BFD,$$

therefore $ABCD \overset{d}{=} ABEF$. In the other case BD and AE will intersect in G, say, (fig. 27). Now one measures off BG on BD and AG on AE and obtains in this way the points G_1, G_2, ... and H_1, H_2, ... Now by the 1st axiom of continuity ([VI 2] or [VI 11]) there is an integer n with the property that G_n coincides with D or on the other hand D lies between G_n and G_{n+1}. Then one draws the parallel lines noted in fig. 27 and reads off immediately from the figure the congruence of triangles and trapezoids corresponding to each other:—

$$\Delta I \equiv \Delta 1, \Delta II \equiv \Delta 2, \ldots, T_1 \equiv T_2, D_1 \equiv D_2,$$

$$\Delta ABG \equiv \Delta ABG.$$

Hence the equivalence by dissection of $ABCD$ and $ABEF$ is proved.

Note that at this point we have used an axiom of continuity. As is well-known from HILBERT [VI 11] one can construct a theory of content which does *not* use an axiom of continuity. Then at this point one must of course be satisfied with the weaker statement that parallelograms with common base and height are *equivalent by completion.*†

Since each triangle is equivalent by dissection to a parallelogram with half the height, it therefore follows by the rule (a): Triangles with *the same base and height are equivalent by dissection.* If one takes in addition the rule (b) one knows immediately that the process (used in school teaching) of " transforming " an n-gon $\mathscr{P}_n$ into an $(n-1)$-gon $\mathscr{P}_{n-1}$ leads to a figure which is not only equal in content but also is *equivalent by dissection*: $\mathscr{P}_n \overset{d}{=} \mathscr{P}_{n-1}$. Also the well-known transformation of a triangle with base AB into a triangle with base AB' leads to a figure equivalent by dissection. Now if one changes two polygons $\mathscr{P}_1$ and $\mathscr{P}_2$ which are equal in content into two triangles $\mathscr{T}_1$ and $\mathscr{T}_2$ with unit bases (and height $h = 2F(\mathscr{P}_1) = 2F(\mathscr{P}_2)$), one has the following relations between the figures

$$\mathscr{P}_\nu \overset{d}{=} \mathscr{T}_\nu (\nu = 1, 2), \quad \mathscr{T}_1 \overset{d}{=} \mathscr{T}_2.$$

Hence it follows by rule (a) that $\mathscr{P}_1 \overset{d}{=} \mathscr{P}_2$.

2. DEHN'S THEOREM

In attempting to transfer these studies to space it is easy to discover that *prisms* which are equal in content are equivalent by dissection (see [VI 2] p. 190). However, in the 19th century people puzzled in vain over the question whether the corresponding statement is true also for pyramids (or for general polyhedra). Thus in 1900 HILBERT posed the following problem to the Mathematical Congress in Paris: *It is required to find two pyramids with the same face as base and equal height which are not equivalent by dissection.*

This problem was solved a few months later by MAX DEHN. He showed that there is a *necessary* condition for the equivalence by dissection of polyhedra. Then it is no longer difficult to find a pair

† Two polygons $\mathscr{P}$ and $\mathscr{Q}$ are " equivalent by completion " ($\mathscr{P} \overset{c}{=} \mathscr{Q}$), if there are finitely many triangles $\mathscr{T}_\nu$ and $\mathscr{T}'_\nu$ congruent in pairs such that

$$\mathscr{P} + \Sigma \mathscr{T}_\nu \overset{d}{=} \mathscr{Q} + \Sigma \mathscr{T}'_\nu.$$

of pyramids with common base and equal height for which Dehn's condition is *not* satisfied.

Theorem 20: *Let $\mathscr{P}$ and $\mathscr{Q}$ be polyhedra equivalent by dissection, having edge angles* † $\alpha_1, \alpha_2, \ldots, \alpha_n$ *and* $\beta_1, \beta_2, \ldots, \beta_m$ *respectively. Then there exist non-negative integers* $\nu_1, \nu_2, \ldots, \nu_n$ *and* $\mu_1, \mu_2, \ldots, \mu_m$ *and an integer* ν *such that*

$$(\nu_1\alpha_1+\nu_2\alpha_2+\ldots+\nu_n\alpha_n)-(\mu_1\beta_1+\mu_2\beta_2+\ldots+\mu_m\beta_m)=\nu\pi. \quad (1)$$

To prove this theorem we want to identify the sub-polyhedra of $\mathscr{P}$ and $\mathscr{Q}$ and to demonstrate that we construct from the " building bricks " $\not{p}_1, \not{p}_2, \ldots, \not{p}_r$ first the polyhedron $\mathscr{P}$ then (from another " building plan ") the polyhedron $\mathscr{Q}$ equivalent by dissection to $\mathscr{P}$. We note the building plan for each of the sub-polyhedra by marking the edges of the different, neighbouring polyhedra (in the 1st or 2nd building plan) on $\not{p}_\rho$. For those who prefer a visual demonstration one can colour the edges of $\not{p}_\rho$ with a suitable dye and this clearly indicates how to build up the polyhedra $\mathscr{P}$ and $\mathscr{Q}$.

In this way a net of lines whose individual lines we will call " boundary lines " is drawn on $\not{p}_\rho$. These boundary lines are

1. edges,
2. segments of edges,
3. lines which are noted on the faces of $\not{p}_\rho$.

An edge angle is now associated uniquely with each of these boundary lines s:—If s is an edge (or part of an edge) of $\not{p}_\rho$ then the edge angle is just the angle included by the two faces intersecting in the edge. We associate the boundary line angle π with a boundary line plotted on a face.

We now construct the sum S of all edge angles and calculate it by two different methods. If we construct from the $\not{p}_\rho$ the polyhedron $\mathscr{P}$, then certain boundary lines will lie on the edge e_τ of $\mathscr{P}$. The sum of these edge angles is an integral multiple of the edge angle α_τ of $\mathscr{P}$. Other boundary lines which belong to edges of $\not{p}_\rho$ can be placed in the interior of $\mathscr{P}$; the sum of the corresponding edge angles is an integral multiple of 2π. The boundary lines lying in the faces of $\not{p}_\rho$ finally yield an edge angle sum which must be a multiple of π. Hence we have for the sum S of all edge angles

$$S=\sum_{N=1}^{n}\nu_N\alpha_N+N_1\pi.$$

† They are measured in radians.

Similarly, the construction of $\mathcal{Q}$ from the sub-polyhedra yields

$$S = \sum_{M=1}^{m} \mu_M \beta_M + N_2 \pi.$$

If one forms the difference of these two representations for S then relation (1) is obtained.

3. Examples on Dehn's Theorem

If the second polyhedron is a right parallelepiped then the necessary condition (1) for the equivalence by dissection is simplified to

$$\sum_{N=1}^{n} \nu_N \alpha_N = N^* \cdot \frac{\pi}{2}. \tag{2}$$

One easily sees that this condition is satisfied for Juel's pyramids. These are pyramids whose base is a square with side of length a and whose vertex is on the normal to the base through the point of intersection of the diagonals of the base (at the height $a/2$). The 8 edge angles of such a pyramid are

$$\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \frac{\pi}{4}, \quad \alpha_5 = \alpha_6 = \alpha_7 = \alpha_8 = \frac{2\pi}{3}. \tag{3}$$

Since one can combine 6 such pyramids to form a cube it follows at once that the edge angles which do not belong to the base are equal to $2\pi/3$. The bases of the pyramids become side faces of the cube. The edges with the angles α_5, α_6, α_7, α_8 lie on the principle diagonals of the cube. At each half principal diagonal three edges of congruent pyramids adjoin and the sum of their angles is 2π; therefore each edge angle must be equal to $2\pi/3$.

Now from (3) it follows that

$$\sum_{\nu=1}^{4} \alpha_\nu + 3 \sum_{\nu=5}^{8} \alpha_\nu = 9\pi.$$

The *necessary* condition (2) for the equivalence by dissection with a cube (and/or a right parallelepiped) is therefore satisfied.† We shall show later (p. 79) that each Juel's pyramid is actually equivalent by dissection to a cube.

First of all, we now give an example of two pyramids with the same base and equal height for which condition (1) is not satisfied and which therefore cannot be equivalent by dissection.

† M. Goldberg [VI 13] has collected further examples of pyramids which are equivalent by dissection to a cube.

Let $\mathscr{J}$ be Juel's pyramid with base edge of length 1 and $\mathscr{P}$ a pyramid with the same base and equal height but with its vertex placed at the height $\frac{1}{2}$ above the midpoint of a side of the square (fig. 28). Let the edge angles α_ν of $\mathscr{P}$ be numbered in the same way as the edges of $\mathscr{P}$ in fig. 28. Then

$$\alpha_1 = \alpha_3 = \frac{\pi}{4}, \quad \alpha_4 = \frac{\pi}{2}.$$

From $\tan \alpha_2 = \frac{1}{2}$ we calculate for later use

$$\sin \alpha_2 = \tfrac{1}{5}\sqrt{5}, \quad \cos \alpha_2 = \tfrac{2}{5}\sqrt{5}. \tag{4}$$

From the solid angle formed by the edges 1, 4 and 5 it follows that

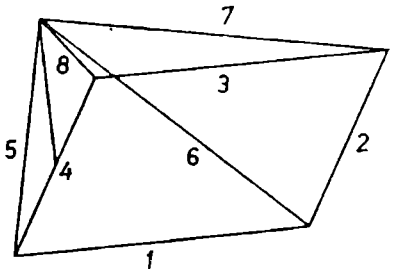


FIG. 28

$\alpha_5 = \alpha_8 = \pi/2$ is also true. Finally we calculate $\alpha_6 = \alpha_7$ from the solid angle with edges 1, 2 and 6. Since the edges 1 and 2 of this angle are perpendicular to each other one finds

$$\cos \alpha_6 = -\cos \alpha_1 \cos \alpha_2 = \frac{-1}{5}\sqrt{10}; \quad \sin \alpha_6 = \frac{1}{5}\sqrt{15}. \tag{5}$$

Now if $\mathscr{P}$ and $\mathscr{J}$ were equivalent by dissection then a relation of the form (1) would have to be satisfied, therefore

$$\nu_1 \cdot \frac{\pi}{4} + \nu_2\alpha_2 + \nu_6\alpha_6 = \nu\pi + \left(\mu_1 \frac{\pi}{4} + \mu_2 \frac{2\pi}{3}\right)$$

or (after multiplication by 12)

$$r\alpha_2 + s\alpha_6 = t\pi$$

where r and s are certain even numbers and t is integral. Hence we must have

$$\sin r\alpha_2 = \pm \sin s\alpha_6. \tag{6}$$

Now †

$$\sin n\alpha = n\cos^{n-1}\alpha\sin\alpha - \binom{n}{3}\cos^{n-3}\alpha\sin^3\alpha + \binom{n}{5}\cos^{n-5}\alpha\sin^5\alpha - + \ldots. \tag{7}$$

If one applies this formula (7) to both sides of (6) then one obtains a contradiction. For by (4), for even r and s, $\sin r\alpha_2$ is rational while $\sin s\alpha_6$ (from (5) and (7)) is of the form $\frac{a}{b}\cdot\sqrt{6}$ with integral a and b. Since $\sqrt{6}$ is not rational, (6) cannot be correct. Our assumption that there exists a relation of the form (1) for $\mathscr{J}$ and $\mathscr{P}$, has therefore led to a contradiction. $\mathscr{P}$ and $\mathscr{J}$ are therefore *not equivalent by dissection.*

Theorem 21: *There are pyramids with the same face as base and equal height which are not equivalent by dissection.*

There are various other examples of this theorem. We want to mention yet another:—The regular tetrahedron ‡ is *not equivalent by dissection* to a tetrahedron $\mathscr{T}_1$ which has the same height and base as $\mathscr{T}$ but whose vertex is vertically above a vertex of the base. The proof of this is left to the reader.

By similar methods one can show that no regular tetrahedron, octahedron, rhombic dodecahedron or icosahedron is equivalent by dissection to a cube [VI 2].

We have still to give the proof of the fact that Juel's pyramids really are equivalent to a cube. For a start it is ascertained only that Dehn's condition (1) is satisfied for them. To do this we require an important theorem by SYDLER [VI 3]:—

Theorem 22: *Let $\mathscr{P}$ be a polyhedron and a_ν ($\nu = 1, 2, \ldots, n$) n arbitrarily given, positive, real numbers which satisfy the equation*

$$a_1 + a_2 + \ldots + a_n = 1.$$

† This formula is obtained most simply from de Moivre's theorem. From

$$(\cos\alpha + i\sin\alpha)^n = \cos n\alpha + i\sin n\alpha = \cos^n\alpha + \binom{n}{1}\cos^{n-1}\alpha\, i\sin\alpha + \ldots$$

formula (7) follows by taking imaginary parts and by taking real parts

$$\cos n\alpha = \cos^n\alpha - \binom{n}{2}\cos^{n-2}\alpha\sin^2\alpha + \binom{n}{4}\cos^{n-4}\alpha\sin^4\alpha - + \ldots. \tag{7'}$$

‡ In the following we call each pyramid with triangular base a tetrahedron. If we mean a triangular pyramid whose side faces are equilateral triangles, we add the adjective " regular ".

Then there are n polyhedra $\mathcal{P}_\nu$ and a polyhedron $\mathcal{R}$ with the following properties:—

1. *The $\mathcal{P}_\nu$ are similar to $\mathcal{P}$ and have the linear similarity factor a_ν,*
2. *$\mathcal{R}$ is equivalent by dissection to a right parallelepiped,*
3. $\mathcal{P} = \mathcal{P}_1 + \mathcal{P}_2 + \ldots + \mathcal{P}_n + \mathcal{R}$.

First of all we assume that $\mathcal{P}$ is a tetrahedron $\mathcal{T}$ (fig. 29). Now let

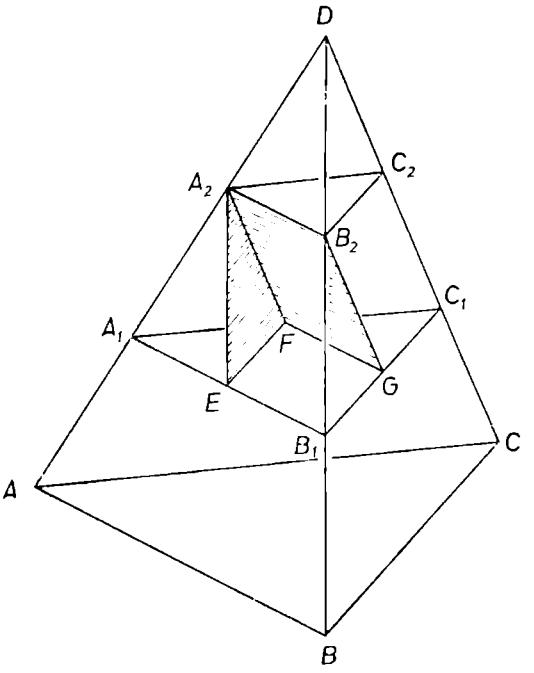


FIG. 29

the edge AD of $\mathcal{P} = ABCD$ be subdivided by the points A_1, A_2, A_3, ..., so that

$$AA_1 : A_1A_2 : A_2A_3 \ldots = a_1 : a_2 : a_3 \ldots$$

is true. Now, through the points A_ν are drawn planes parallel to ABC and cutting the triangles $A_\nu B_\nu C_\nu$ on the tetrahedron. Next we draw through A_2 the plane parallel to the plane BCD and intersecting A_1B_1 and A_1C_1 in the points E and F (fig. 29). Finally, let FG be parallel to EB_1 parallel to A_2B_2.

Thus the frustum $A_1B_1C_1A_2B_2C_2$ of the pyramid is decomposed into a tetrahedron ($\mathcal{T}_2 = A_1A_2EF$) and two prisms ($EB_1GFA_2B_2$ and $A_2B_2FGC_2C_1$). The tetrahedron $\mathcal{T}_2$ is similar to the whole tetrahedron $\mathcal{T}$ (and to the tetrahedron $\mathcal{T}_n = DA_{n-1}B_{n-1}C_{n-1}$ at the vertex) and has similarity factor a_2. Now, if we undertake a

corresponding decomposition in the other parallel layers, we have for $\mathscr{T}$ a decomposition of the form

$$\mathscr{T} \overset{d}{=} t_1 + t_2 + \ldots + t_n + \mathscr{R}, \tag{8}$$

where t_ν is a tetrahedron similar to $\mathscr{T}$ with similarity factor a_ν and the sum of all the prisms arising in this decomposition is combined to form $\mathscr{R}$. Since each prism is equivalent by dissection to a rectangle, this is also true for $\mathscr{R}$. If, now, $\mathscr{P}$ is any polyhedron, we decompose it by suitable cuts into tetrahedra.† Then we have

$$\mathscr{P} \overset{d}{=} \mathscr{T}_1 + \mathscr{T}_2 + \ldots + \mathscr{T}_N,$$

and from (8)

$$\mathscr{T}_\sigma \overset{d}{=} t_{\sigma_1} + t_{\sigma_2} + \ldots + t_{\sigma_n} + \mathscr{R}_\sigma.$$

Now one can again form polyhedra p_ν similar to $\mathscr{P}$ from the tetrahedra similar to $\mathscr{T}_\sigma$ and hence it follows that

$$\mathscr{P} \overset{d}{=} \sum_{\sigma=1}^{N} \mathscr{T}_\sigma = \sum_{\nu=1}^{n} p_\nu + \mathscr{R}.$$

Now, from Theorem 22, the equivalence by dissection of Juel's pyramid and the right parallelepiped follows immediately: If $\mathscr{J}$ is a Juel's pyramid and $a_\nu = \frac{1}{6}$ then we obtain

$$\mathscr{J} = 6j + \mathscr{R},$$

where j is a pyramid similar to $\mathscr{J}$ with linear similarity factor $\frac{1}{6}$. However one can now combine the 6 pyramids j into a cube and hence

$$\mathscr{J} \overset{d}{=} W + \mathscr{R} \overset{d}{=} \mathscr{R}_1.$$

One can also prove the equivalence by dissection directly by suitable cuts in Juel's pyramid.

The following theorem presents another application of Theorem 22.

Theorem 23: *Polyhedra which are equivalent by completion are equivalent by dissection.*

The corresponding theorem in the plane for polygons is a simple consequence of the fact that polygons which are equivalent by completion are equal in content and polygons with equal content are also always equivalent by dissection (p. 71). In space this

† See e.g. [VI 2] on the decomposition of polyhedra to tetrahedra.

simple argument is not possible. We prove Theorem 23 as follows: Let

$$\mathscr{A}+\mathscr{C}\stackrel{d}{=}\mathscr{B}+\mathscr{D},\quad \mathscr{C}\stackrel{d}{=}\mathscr{D},\quad V(\mathscr{A})=V(\mathscr{B})=V_1, \tag{9}$$

$$V(\mathscr{C})=V(\mathscr{D})=V_2.$$

Then we choose a natural number n for which

$$n^2-1>\frac{3V_2}{V_1}. \tag{10}$$

Now, by Theorem 22, $\mathscr{A}$ is decomposed into n similar (and congruent to each other) polyhedra $\mathscr{A}'$ and a residual polyhedron $\mathscr{R}$ (equivalent by dissection to a right parallelepiped):—

$$\mathscr{A}=n\mathscr{A}'+\mathscr{R}, \tag{11}$$

where

$$V(\mathscr{A}')=\frac{1}{n^3}V_1=V_1'. \tag{12}$$

From the residual polyhedron $\mathscr{R}$ are now cut off n polyhedra $\mathscr{C}'$ which are similar to $\mathscr{C}$ and have volume

$$V_2'=\frac{1}{n^3}V_2. \tag{13}$$

Of course, it must first be shown that this is always possible. First of all let $\mathscr{C}$ be a tetrahedron $\mathscr{T}=ABCD$. Then $\mathscr{T}$ is part of a prism $\mathscr{P}^*=ABCDEF$ which one can construct upon 3 of the 4 edges of $\mathscr{T}$. Then we have $V(\mathscr{P}^*)=3V(\mathscr{T})$. The required separation of n similar tetrahedra $\mathscr{T}'$ from $\mathscr{R}$ can now be carried out as follows:—From $\mathscr{R}$ one removes n cubes of volume $3V_2'$, converts them into prisms similar to $\mathscr{P}^*$ and cuts from these the n tetrahedra $\mathscr{T}'$ similar to $\mathscr{T}$. This procedure is always possible if the volume of $\mathscr{R}$ is greater than that of the n prisms and this follows from our choice of the number n. We have

$$V(\mathscr{R})=V_1-nV_1'=V_1\left(1-\frac{1}{n^2}\right)$$

from (11), (9) and (12) and

$$3nV_2'=\frac{3}{n_2}V_2$$

from (13). Hence, from (10), it follows that

$$V(\mathscr{R})=V_1\left(1-\frac{1}{n^2}\right)>\frac{3}{n^2}V_2=3nV_2'.$$

If $\mathscr{C}$ is any polyhedron then one decomposes it by suitable cuts into tetrahedra, carries out the decomposition of the individual tetrahedra and then combines the sub-tetrahedra again into polyhedra $\mathscr{C}'$ which are similar to $\mathscr{C}$. Then we can represent the described partitioning of $\mathscr{R}$ as follows:—

$$\mathscr{R} \overset{d}{=} n\mathscr{C}' + \mathscr{S}. \tag{14}$$

Then we have from (11) and (14)

$$\mathscr{A} \overset{d}{=} n\mathscr{A}' + n\mathscr{C}' + \mathscr{S} \overset{d}{=} n(\mathscr{A}' + \mathscr{C}') + \mathscr{S}. \tag{15}$$

From (9) it follows for the polyhedra $\mathscr{A}'$ and $\mathscr{C}'$ similar to $\mathscr{A}$ and $\mathscr{C}$ that

$$\mathscr{A}' + \mathscr{C}' \overset{d}{=} \mathscr{B}' + \mathscr{D}'. \tag{16}$$

Since

$$V(\mathscr{B}') = V(\mathscr{A}') = V_1', \quad V(\mathscr{D}') = V(\mathscr{C}') = V_2'$$

from (15) and (16) we get

$$\mathscr{A} \overset{d}{=} n(\mathscr{B}' + \mathscr{D}') + \mathscr{S} \overset{d}{=} n\mathscr{B}' + (n\mathscr{C}' + \mathscr{S}) \overset{d}{=} n\mathscr{B}' + \mathscr{R}. \tag{17}$$

If one applies to $\mathscr{B}$ the same decomposition as applied to $\mathscr{A}$ (from Theorem 22), one obtains

$$\mathscr{B} = n\mathscr{B}' + \mathscr{R}'. \tag{18}$$

From (9), (16) and (18) it follows that $\mathscr{R}$ and $\mathscr{R}'$ have equal volumes. But right parallelepipeds with equal volumes are equivalent by dissection. Thus $\mathscr{R} \overset{d}{=} \mathscr{R}'$ and from (17) and (18)

$$\mathscr{A} \overset{d}{=} n\mathscr{B}' + \mathscr{R} \overset{d}{=} \mathscr{B}$$

is true.

We will now use both Theorems 22 and 23 to derive *necessary and sufficient* conditions, first obtained by HADWIGER, for the equivalence by dissection of polyhedra.

For over fifty years it has been known that a Dehn's condition of the form (1) must be satisfied if two polyhedra are equivalent by dissection. Whether this condition is also *sufficient* is so far still not known. Now Hadwiger's condition which we derive (Theorem 26, p. 88) is indeed *necessary and sufficient* but, unfortunately, it cannot be used to decide whether a given pair of polyhedra of equal content are really equivalent by dissection. The reason for this is that, in the formulation of this condition, we use a set-theoretical existence statement. Nevertheless, this theorem of Hadwiger is an important contribution to the theory of content. For its derivation we develop a kind of " algebra of polyhedra ".

4. The Algebra of Polyhedra

In the following we denote by $\lambda\mathscr{P}$ a polyhedron similar to $\mathscr{P}$ with *linear* similarity factor λ. Therefore for the volume $V(\mathscr{P})$

$$V(\lambda\mathscr{P}) = \lambda^3 \,.\, V(\mathscr{P}) \tag{19}$$

is true. In general

$$(\lambda+\mu)\mathscr{P} \neq \lambda\mathscr{P}+\mu\mathscr{P}.$$

This can be seen, e.g. for $\lambda+\mu = 1$ using Theorem 22: thence

$$1\,.\,\mathscr{P} = \mathscr{P}_1+\mathscr{P}_2+\mathscr{R} = \lambda\mathscr{P}+\mu\mathscr{P}+\mathscr{R}.$$

We proceed to develop an algebra of polyhedra for which the distributive law holds and also multiplication by negative numbers is meaningful. First of all we define multiplication by 0 by means of the " empty " polyhedron $\mathscr{O}$:—

$$0\mathscr{P} = \mathscr{O}, \quad \mathscr{P}+\mathscr{O} = \mathscr{P}.$$

Since each right parallelepiped is equivalent by dissection to a cube, we can also formulate Sydler's theorem (Theorem 22) as follows:

Theorem 22a: *Let* $a_\nu > 0$ *and* $\sum_{\nu=1}^{n} a_\nu = 1$. *Then it is true for each polyhedron* $\mathscr{P}$ *that*

$$\mathscr{P} \overset{d}{=} \lambda\mathscr{U} + \sum_{\nu=1}^{n} a_\nu\mathscr{P} \tag{20}$$

where $\mathscr{U}$ *is the unit cube,* $\lambda > 0$ *and*

$$\lambda^3 = V(\mathscr{P})\left(1 - \sum_{\nu=1}^{n} a_\nu^3\right).$$

Next we state the definition:—*Two polyhedra* $\mathscr{P}$ *and* $\mathscr{Q}$ *are said to be equivalent modulo* $\mathscr{U}$,

$$\mathscr{P} \approx \mathscr{Q} \bmod \mathscr{U} \text{ (or simply } \mathscr{P} \approx \mathscr{Q}),$$

if there exist two numbers α *and* β *for which the relation*

$$\mathscr{P}+\alpha\mathscr{U} \overset{d}{=} \mathscr{Q}+\beta\mathscr{U}, \quad \alpha \geqq 0, \quad \beta \geqq 0, \quad \alpha\,.\,\beta = 0,$$

exists. In other words:—Two polyhedra are considered to be equivalent if the one with the larger volume can be made up from parts of the other and of a right parallelepiped. With this new notation (20) assumes the simple form:—

$$\mathscr{P} \approx \sum_{\nu=1}^{n} a_\nu\mathscr{P}. \tag{20'}$$

There are the following rules for equivalence of polyhedra:

1. $\mathcal{P} \approx \mathcal{P}$.
2. From $\mathcal{P} \approx \mathcal{Q}$ it follows that $\mathcal{Q} \approx \mathcal{P}$.
3. From $\mathcal{P} \approx \mathcal{Q}$ and $\mathcal{Q} \approx \mathcal{R}$ it follows that $\mathcal{P} \approx \mathcal{R}$.
4. From $\mathcal{P} \approx \mathcal{Q}$ and $\mathcal{P}+\mathcal{R} \approx \mathcal{Q}+\mathcal{S}$, it follows that $\mathcal{R} \approx \mathcal{S}$.

These rules are proved directly from the definition. We will only prove rule 3. Let

$$\mathcal{P}+\alpha\mathcal{U} \stackrel{d}{=} \mathcal{Q}+\beta\mathcal{U}, \quad \mathcal{Q}+\gamma\mathcal{U} \stackrel{d}{=} \mathcal{R}+\delta\mathcal{U}$$

and $\beta \geqq \gamma$, say. Then

$$\beta\mathcal{U} \stackrel{d}{=} \gamma\mathcal{U}+\xi\mathcal{U} \text{ (with } \beta^3 = \gamma^3+\xi^3),$$

therefore

$$\mathcal{P}+\alpha\mathcal{U} \stackrel{d}{=} \mathcal{Q}+\gamma\mathcal{U}+\xi\mathcal{U} \stackrel{d}{=} \mathcal{R}+\delta\mathcal{U}+\xi\mathcal{U}$$

or

$$\mathcal{P}+\alpha\mathcal{U} \stackrel{d}{=} \mathcal{R}+\eta\mathcal{U}, \quad (\eta^3 = \delta^3+\xi^3).$$

However, this means that $\mathcal{P} \approx \mathcal{R}$.

For each right parallelepiped $\mathcal{A}$ it is true that $\mathcal{A} \approx \mathcal{O}$ for

$$\mathcal{A}+0\mathcal{U} = \mathcal{O}+\lambda\mathcal{U}.$$

From Dehn's theorem (Theorem 22) we have the result that there exist polyhedra which are not equivalent by dissection to a cube. We can formulate this statement in the new terminology as follows:

Theorem 20a: *There exist polyhedra $\mathcal{P}$ which are not equivalent to the empty polyhedron:—*

$$\mathcal{P} \not\approx \mathcal{O}.$$

In order to have negative factors also in the algebra of polyhedra for use in other investigations, we introduce the concept of the *negative polyhedron* by the rule

$$-\mathcal{P} = \lambda\mathcal{U}-\mathcal{P}^\circ, \tag{21}$$

where $\lambda\mathcal{U}$ is a sufficiently large cube which covers $\mathcal{P}$ so that each point of $\mathcal{P}$ is an interior point of the cube and $\mathcal{P}^\circ$ denotes the open kernel of $\mathcal{P}$. In this way the usual subtraction of sets in (21) means that *$-\mathcal{P}$ is the closed set of points which is obtained from the cube $\lambda\mathcal{U}$ when the interior points of $\mathcal{P}$ are removed.*

Of course $-\mathcal{P}$ is not uniquely determined by $\mathcal{P}$. If one chooses another cube $\lambda'\mathcal{U}'$ (whose faces need not be parallel to those of $\lambda\mathcal{U}$) one obtains another negative polyhedron $(-\mathcal{P})'$. However,

the two negative polyhedra $(-\mathscr{P})$ and $(-\mathscr{P})'$ belonging to $\mathscr{P}$ are equivalent. For, from

$$\lambda\mathscr{U}-\mathscr{P}^{\circ}=\mathscr{Q}_1,\quad \lambda'\mathscr{U}'-\mathscr{P}^{\circ}=\mathscr{Q}_2,\quad \lambda\mathscr{U}\approx\lambda'\mathscr{U}'\approx\mathscr{O},$$

therefore

$$\mathscr{Q}_1+\mathscr{P}^{\circ}\approx\mathscr{Q}_2+\mathscr{P}^{\circ},$$

it follows from rules 4 and 1, that

$$\mathscr{Q}_1\approx\mathscr{Q}_2.$$

Hence in our algebra of polyhedra negative scalar factors $\alpha, \beta, \ldots$ are also possible: $(-\alpha)\mathscr{P}$ means simply the negative polyhedron $-(\alpha\mathscr{P})$. Now we note the rules arising from this definition and the earlier statements:—

5. $\mathscr{Q}+(-\mathscr{Q})\approx\mathscr{O}$.

6. The statements

$$\mathscr{P}+\mathscr{Q}\approx\mathscr{R} \text{ and } \mathscr{P}\approx\mathscr{R}-\mathscr{Q}=\mathscr{R}+(-\mathscr{Q})$$

are equivalent.

7. $\alpha(\mathscr{P}+\mathscr{Q})\approx\alpha\mathscr{P}+\alpha\mathscr{Q}$.

8. $\alpha(\beta\mathscr{P})\approx(\alpha\beta)\mathscr{P}$.

9. $\alpha\mathscr{P}+\beta\mathscr{P}\approx(\alpha+\beta)\mathscr{P}$.

We omit the simple proofs for rules 5 to 8 and limit ourselves to the proof of 9. For *positive* numbers α and β this rule follows from Theorem 22a. For, according to it,

$$\mathscr{P}\approx\frac{\alpha}{\alpha+\beta}\mathscr{P}+\frac{\beta}{\alpha+\beta}\mathscr{P}.$$

Now let $\alpha>\beta>0$ be true and $\alpha_1=\alpha$, $\beta_1=-\beta$. Then from Theorem 22a

$$(\alpha-\beta)\mathscr{P}+\beta\mathscr{P}\approx\alpha\mathscr{P}.$$

Hence, from rule 6, it follows that

$$(\alpha-\beta)\mathscr{P}\approx\alpha\mathscr{P}-\beta\mathscr{P} \text{ or } (\alpha_1+\beta_1)\mathscr{P}\approx\alpha_1\mathscr{P}+\beta_1\mathscr{P}.$$

If both coefficients are negative, one uses the fact that $\mathscr{P}\approx\mathscr{Q}$ implies the equivalence $-\mathscr{P}\approx-\mathscr{Q}$.

5. The Basis Polyhedra

To derive his theorems on equivalence by dissection HADWIGER [VI 6] uses set-theoretical methods which Hamel used in similar form for real numbers.†

† See e.g. [VI 5] p. 135.

Using the axiom of choice one proves in set theory ([VI 5], p. 130):—*Each set can be well-ordered.* As is well-known a set is said to be *well-ordered, if an ordering is defined in it so that each non-empty subset has a first element.*

The rational numbers between 0 and 1 can be well-ordered by the usual enumeration by increasing denominators (for equal denominators by increasing numerators):—

$$\frac{1}{2},\ \frac{1}{3},\ \frac{2}{3},\ \frac{1}{4},\ \frac{3}{4},\ \frac{1}{5},\ \frac{2}{5},\ \frac{3}{5},\ \frac{4}{5},\ \ldots. \tag{22}$$

If we write $r_1 \angle r_2$ when r_2 in the ordering (22) occurs to the right of r_1 then the set of rational numbers is well-ordered by *this* ordering. If one takes the usual ordering as basis, instead of this, then each non-empty subset of the set of rational numbers does not necessarily have a first element.

Therefore one can really state the well-ordering for the rational numbers. For the set of real numbers between 0 and 1 (and more so for complex sets) the well-ordering theorem is purely a statement of existence.†

Now let $\{\mathscr{P}\}$ be the class (non-empty by Theorem 20a) of all polyhedra which are not equivalent by dissection to any right parallelepiped: $\mathscr{P} \approx \mathscr{O}$. Let this set be well-ordered

$$\mathscr{P}_{\kappa_0} \angle \mathscr{P}_{\kappa_1} \angle \mathscr{P}_{\kappa_2} \angle \ldots. \tag{23}$$

We now say that a polyhedron $\mathscr{P} \in \{\mathscr{P}\}$ is " pre-dependent " if there exists an equivalence relation of the form

$$\mathscr{P} \approx \Sigma \alpha_\kappa \mathscr{P}_\kappa \tag{24}$$

where in the (finite) sum only $\mathscr{P}_\kappa$ for which $\mathscr{P}_\kappa \angle \mathscr{P}$ is true, are to appear.

Next let $\{\mathscr{A}\}$ be the class of polyhedra which are *not pre-dependent.* Certainly $\mathscr{P}_{\kappa_0}$ belongs to it and therefore $\{\mathscr{A}\}$ is not empty. $\{\mathscr{A}\}$, as a subset of a well-ordered set, is itself well-ordered with respect to the same ordering relation. These polyhedra belonging to $\{\mathscr{A}\}$ are called (in analogy to Hamel's " basis numbers ") *basis polyhedra.*

Theorem 24: *The basis polyhedra are linearly independent, i.e. from*

$$\sum_{\nu=1}^{n} \alpha_{\tau_\nu} \mathscr{A}_{\tau_\nu} \approx \mathscr{O} \tag{25}$$

† In an intuitionistic theory (see e.g. [VI 10]) non-constructive existence statements are inadmissible. Under such hypotheses the conclusions of HAMEL and the statements made in this section for content theory are not valid.

it follows that

$$\alpha_{\tau_1} = \alpha_{\tau_2} = \ldots = \alpha_{\tau_n} = 0.$$

To prove this we can assume that $n \geqq 2$. For $\mathscr{A}_{\tau_1} \approx \mathcal{O}$ would certainly contradict the definition of the basis polyhedra: $\{\mathscr{A}\}$ is, of course, a subset of $\{\mathscr{P}\}$. However, for $n \geqq 2$ and $\alpha_{\tau_n} \neq 0$, it would follow from (25) that

$$\mathscr{A}_{\tau_n} \approx \sum_{\nu=1}^{n-1} \lambda_\nu \mathscr{A}_{\tau_\nu}, \quad \lambda_\nu = -\frac{\alpha_{\tau_\nu}}{\alpha_{\tau_n}}.$$

But then $\mathscr{A}_{\tau_n}$ would be, contrary to the hypothesis, pre-dependent.

Theorem 25: *Each polyhedron $\mathscr{P} \in \{\mathscr{P}\}$ has a unique representation as a finite † sum of the form*

$$\mathscr{P} \approx \Sigma \alpha_\tau \mathscr{A}_\tau \tag{26}$$

with (positive or negative) real coefficients α_τ.

We show first of all that there exists *at least* one such representation. The proof is carried through indirectly:—Let $\{\mathscr{S}\}$ be the (non-empty) subset of $\{\mathscr{P}\}$ for which a relation (26) does not hold. As a subset of a well-ordered set $\{\mathscr{S}\}$ is well-ordered. Let the first element of this set be $\mathscr{S}_1$. If $\mathscr{S}_1$ were a basis polyhedron then we would have $\mathscr{S}_1 = \mathscr{A}_{\tau_\rho}$, therefore also $\mathscr{S}_1 \approx \mathscr{A}_{\tau_\rho}$ and hence a representation of the form (26). Therefore we must suppose that $\mathscr{S}_1$ does not belong to $\{\mathscr{A}\}$. However, $\mathscr{S}_1$ is then pre-dependent, i.e. there is a representation of the form

$$\mathscr{S}_1 \approx \sum_\kappa \beta_\kappa \mathscr{P}_\kappa \tag{27}$$

where $\mathscr{P}_\kappa \angle \mathscr{S}_1$. Since $\mathscr{S}_1$ is the first element in our well-ordering for which a representation (26) does not hold, the $\mathscr{P}_\kappa$ appearing in (27) are all representable as follows

$$\mathscr{P}_\kappa \approx \sum_\rho \alpha_{\kappa\rho} \mathscr{A}_\rho, \quad \mathscr{A}_\rho \angle \mathscr{P}_\kappa.$$

Hence it follows that

$$\mathscr{S}_1 \approx \sum_\kappa \sum_\rho \beta_\kappa \alpha_{\kappa\rho} \mathscr{A}_\rho \approx \sum_\rho \gamma_\rho \mathscr{A}_\rho, \quad \mathscr{A}_\rho \angle \mathscr{P}_\kappa \angle \mathscr{S}_1$$

where

$$\gamma_\rho = \sum_\kappa \beta_\kappa \alpha_{\kappa\rho}.$$

Thus we would have obtained for $\mathscr{S}_1$ a representation of the form (26).

† All the sums appearing in this section are finite.

Now let there exist two representations of this kind for $\mathscr{P}$. Then one would have, say,

$$\mathscr{P} \approx \sum_{\tau=1}^{n} \alpha_\tau \mathscr{A}_\tau \approx \sum_{\tau=1}^{n} \beta_\tau \mathscr{A}_\tau,$$

where *some* α_τ or β_τ respectively, could vanish. Hence it would now follow that

$$\sum_{\tau=1}^{n} (\alpha_\tau - \beta_\tau)\mathscr{A}_\tau \approx \mathscr{O}$$

in contradiction to Theorem 24. Thus Theorem 25 is completely proved. From the definition of equivalence it is possible to replace the representation (26) by another in which only positive coefficients appear:—

Theorem 25a: *There exists a class* $\{\mathscr{A}_\tau\}$ *of basis polyhedra such that to each polyhedron* $\mathscr{P} \in \{\mathscr{P}\}$ *can be related two sets of non-negative (uniquely determined) coefficients* (ξ, β_τ) *and* (η, γ_τ) $(\xi \,.\, \eta = 0,$ $\beta_\tau \,.\, \gamma_\tau = 0)$ *for which the equivalence by dissection*

$$\mathscr{P} + \eta\mathscr{U} + \Sigma\gamma_\tau\mathscr{A}_\tau \overset{d}{=} \xi\mathscr{U} + \Sigma\beta_\tau\mathscr{A}_\tau \tag{28}$$

is true.

For this one only needs to write (26) so that the positive coefficients $\alpha_\tau = \beta_\tau$ are on the right and the negative $\alpha_\tau = -\gamma_\tau$ on the left of the equivalence sign:†

$$\mathscr{P} + \Sigma\gamma_\tau\mathscr{A}_\tau \approx \Sigma\beta_\tau\mathscr{A}_\tau.$$

Hence, immediately from the definition of equivalence one obtains (28) and there remains only the uniqueness of the coefficients ξ and η to be demonstrated. However, this follows from the equation for the volume

$$V(\mathscr{P}) = \xi^3 - \eta^3 + \Sigma(\beta_\tau^3 - \gamma_\tau^3)V(\mathscr{A}_\tau).$$

Then, since $\xi \,.\, \eta = 0$, ξ and η are uniquely determined.

The coefficients α_τ in (26) are functionals of the polyhedra $\mathscr{P}$. To indicate this relation we write

$$\alpha_\tau = \chi_\tau(\mathscr{P}). \tag{29}$$

Each basis polyhedron $\mathscr{A}_\tau \in \{\mathscr{A}\}$ gives rise to a functional $\chi_\tau(\mathscr{P})$. If $\{\mathscr{A}_\tau\}$ does not appear in the representation (26) for $\mathscr{P}$ then $\chi_\tau(\mathscr{P}) = 0$. For each $\mathscr{P}$ there are, from Theorem 25, only finitely many functionals $\chi_\tau(\mathscr{P})$ different from 0. Now we can formulate

† Note that for each number τ at least one of the two numbers γ_τ, β_τ vanishes.

the criterion which we mentioned earlier for the equivalence by dissection of polyhedra as follows:—

Theorem 26: *Two polyhedra are equivalent by dissection if and only if*

$$V(\mathcal{P}) = V(\mathcal{Q}), \quad \chi_\tau(\mathcal{P}) = \chi_\tau(\mathcal{Q}) \tag{30}$$

is true for all † τ.

The *necessity* of condition (30) follows immediately from Theorem 24 and the fact that polyhedra which are equivalent by dissection are equivalent. That (30) is also *sufficient* for the equivalence by dissection follows from Theorems 25a and 23:—From

$$\mathcal{P} \approx \Sigma \alpha_\tau \mathcal{A}_\tau, \quad \mathcal{Q} \approx \Sigma \alpha_\tau \mathcal{A}_\tau$$

we obtain from Theorem 25a the equivalence by dissection

$$\mathcal{P} + \eta \mathcal{U} + \Sigma \gamma_\tau \mathcal{A}_\tau \stackrel{d}{=} \mathcal{Q} + \eta \mathcal{U} + \Sigma \gamma_\tau \mathcal{A}_\tau.$$

This means, however, that $\mathcal{P}$ and $\mathcal{Q}$ are equivalent by completion. Hence, from Theorem 23, it follows that $\mathcal{P} \stackrel{d}{=} \mathcal{Q}$.

The criterion expressed in Theorem 26 has the satisfactory property of being necessary and sufficient. However, it cannot be used in any way to decide for two polyhedra, which are given equal in content and with their geometrical configuration (edges, angles, etc.) known, whether or not they are equivalent by dissection. Because the set of basis polyhedra is not known, one does not know at once whether $\{\mathcal{A}\}$ is denumerable or not. Nevertheless Hadwiger's theorem (Theorem 26) is of great significance in the theory of content. From it we obtain important statements about the cardinality of sets of polyhedra which are equal in content but not equivalent by dissection, and also about the theory of functionals of polyhedra. From Theorem 26 there follows at once

Theorem 26a: *There is a system of polyhedra which are equal in content with the cardinality of the continuum and these are not equivalent by dissection in pairs.*

Now let c be an arbitrary, positive, real number. Then, in the representation

$$\mathcal{P} = \sum_{\tau=1}^{n} \alpha_\tau \mathcal{A}_\tau \quad (n \geqq 2)$$

† We write in our representations in which, of course, always only finite sums appear, α_τ or $\chi_\tau(\mathcal{P})$ with an integral index τ. However this does not mean that the set of basis polyhedra is denumerable.

we choose the numbers α_τ † so that

$$V(\mathscr{P}) = \sum_{\tau=1}^{n} \alpha_\tau^3 V(\mathscr{A}_\tau) = c.$$

For the α_τ there are as many possibilities as the cardinal number of the continuum. All the polyhedra obtained in this way for different series of numbers α_τ ($\tau = 1, 2, 3, \ldots, n$) are equal in content but not equivalent by dissection, since from

$$\sum_{\tau=1}^{n} \alpha'_\tau \mathscr{A}_\tau \approx \sum_{\tau=1}^{n} \alpha''_\tau \mathscr{A}_\tau$$

it would follow (Theorem 24) that $\alpha'_\tau = \alpha''_\tau$. It might be noted that Theorem 26a has also been derived by SYDLER [VI 3] directly from Theorem 22.

Another interesting consequence can be drawn for the theory of functionals of polyhedra from Hadwiger's theorem. The functional $\phi(\mathscr{P}) = \chi_\tau(\mathscr{P})$ defined in (29) has, as one sees immediately, the following properties:—

1. $\phi(\mathscr{P}) = \phi(\mathscr{Q})$ when $\mathscr{P} \overset{d}{=} \mathscr{Q}$.
2. $\phi(\mathscr{P}+\mathscr{Q}) = \phi(\mathscr{P}) + \phi(\mathscr{Q})$.
3. $\phi(\lambda\mathscr{P})$ is a continuous function of λ for $\lambda > 0$.

The last property follows from $\phi(\lambda\mathscr{P}) = \lambda\phi(\mathscr{P})$.

Another functional has these three properties, namely the volume $V(\mathscr{P})$. Now for the volume two other properties, which $\chi_\tau(\mathscr{P})$ does not have, can be added:—

4. $V(\mathscr{U}) = 1$ for the unit cube,
5. $V(\mathscr{P}) \geqq 0$.

Property 4 is not particularly important since one can always show, using a suitable numerical factor λ_τ, that $\chi_\tau^*(\mathscr{P}) = \lambda_\tau \chi_\tau(\mathscr{P})$ has this property. On the other hand 5 is an *important characteristic of the volume.* One can show that there is exactly one functional $\phi(\mathscr{P})$ which has the properties 1, 2, 4 and 5, namely the volume $V(\mathscr{P})$ [VI 9]. In this way one obtains a particularly elegant basis of elementary content theory.

We shall now show that the hypothesis $\phi(\mathscr{P}) \geqq 0$ for the volume is quite characteristic:—If one does without this condition then

† We limit ourselves to numbers $\alpha_\tau \geqq 0$.

there exists, besides the volume, the functional $\chi_\tau(\mathscr{P})$ which has the properties 1, 2 and 3. HADWIGER showed further [VI 6], that each functional $\phi(\mathscr{P})$ with the properties 1, 2 and 3 can be written in the form

$$\phi(\mathscr{P}) = cV(\mathscr{P}) + \Sigma c_\tau \chi_\tau(\mathscr{P}).$$

That means that $V(\mathscr{P})$ and the coefficients $\chi_\tau(\mathscr{P})$ are essentially the only functionals with the properties 1, 2 and 3.

Thus all these properties are important: the problem of finding a sufficient criterion which can be applied for the equivalence by dissection of polyhedra which are equal in content has not yet been solved.

However, SYDLER [VI 7, 8] achieved a noteworthy step in this direction by a complicated decomposition process. He showed the following: If Dehn's condition for two polyhedra is satisfied, their difference is equivalent to a polyhedron whose dihedral angles are multiples of $\pi/4$. If one could manage to show also that such polyhedra are equivalent to a cube, one would have proved that the necessary condition (1) (together with equality of the volume) is also sufficient.

Chapter VII

THE DECOMPOSITION OF RECTANGLES INTO INCONGRUENT SQUARES

1. The Problem

In the 19th and 20th centuries attempts to demonstrate the unsolvability of some problems with which mathematicians had struggled for many centuries, were successful. Among these problems were the squaring of the circle (Ch. X), the parallel postulate and the trisection of the angle (Ch. XI). The idea that the preceding problems might perhaps be unsolvable evolved only after generations of mathematicians had sought in vain for the solutions.

Sometimes it is in exactly the opposite way that one comes to a new understanding. When the question emerged whether it might be possible to decompose a square into incongruent sub-squares, many mathematicians were of the opinion that this problem was not solvable. So it was reported in 1930 by Kraitchik [VII 2] (p. 272):

> " ...il n'est pas possible de décomposer un carré donné en un nombre fini de carrés inégaux deux à deux. Cette proposition quoique non démontrée paraît être vraie. Elle nous a été communiquée par M. Lusin, professeur à Moscou."

R. Sprague also tackled the problem, first of all with the aim of proving its unsolvability. But, he then found † a decomposition of the square into 55 different sub-squares [VII 4]. A year later an extensive work by R. L. Brooks, C. A. B. Smith, A. H. Stone and W. T. Tutte was published and this contained *inter alia* the decomposition into only 26 incongruent sub-squares [VII 7].

† In the literature it is sometimes reported (e.g. [VII 10]) that the work [VII 7] gave the first decomposition of a square into incongruent sub-squares. This is not true, however. According to [VII 10] the first decomposition of a *rectangle* into incongruent sub-squares was found in 1925. The book [VII 2] by Kraitchik, 1930, contains as an example only one decomposition of this kind in which there appear pairs of equal squares and also two equal squares with a side in common.

In the mathematical journals dissection problems of the kind discussed here are usually listed under " recreational mathematics ". But it is worth noting that our problem has a technical significance which is not completely obvious at first sight. In the works [VII 7] and [VII 10] the mathematical problem is interpreted immediately in terms of electrical engineering and in the proofs there is a mixture of mathematical and physical arguments. In this chapter we tackle the problem first of all in purely mathematical terms and only at the end discuss the possibility of a technical interpretation.

We start with some conventions for a suitable notation. Fig. 30 shows the decomposition of a rectangle with sides in the ratio 33:32 into 9 incongruent squares and fig. 31 that of a square into 24 incongruent squares. Since one cannot very well print diagrams for all the decompositions under discussion, a simple and uniform lettering for any decomposition of this kind is desired. From a suggestion of C. J. BOUWKAMP [VII 8] this can be done as follows:— Rectangles are orientated so that the greater side lies horizontally and the greatest of the four corner squares is placed in the upper left-hand corner (fig. 30). Then in the decomposition figure we call the lines which are made up of sides of squares lying in a horizontal line, *horizontal cuts*. In fig. 30 these are the lines $A_\nu B_\nu$ ($\nu = 1, 2, \ldots, 6$). We now make a note of the decompositions by recording the lengths of the sides of all the sub-squares, where we begin with the squares lying below the first horizontal cut. They are listed going from left to right. This is followed by the squares (always collected in brackets) lying below the subsequent cuts. Our code for fig. 30 is therefore

$$(18,15)\quad (7,8)\quad (14,4)\quad (10,1)\quad (9). \tag{1}$$

Fig. 31 is to be coded as follows

$$\begin{gathered}(55,39,81)\quad (16,9,14)\quad (4,5)\quad (3,1)\quad (20)\quad (56,18)\\ (38)\quad (30,51)\quad (64,31,29)\quad (8,43)\quad (2,35)\quad (33).\end{gathered} \tag{2}$$

Then the following code (3) denotes a decomposition of the square into 38 different sub-squares given by R. L. BROOKS [VII 7]:—

$$\begin{gathered}(2132,1440,1348)\quad (256,1092)\quad (692,584,164)\quad (420)\\ (108,281,615)\quad (840,758,534,627,173)\quad (454)\quad (199,893)\\ (120,694)\quad (244,310)\quad (217,984)\quad (82,900)\quad (922)\quad (527)\\ (1587)\quad (287,240)\quad (7,1177)\quad (104,1130)\quad (1026).\end{gathered} \tag{3}$$

One should note the difference between the decompositions of figs. 30 and 31:—In fig. 31 one can combine a subset of the sub-squares to form a rectangle. In fig. 30 this is not possible. The square of fig. 31 contains (upper left) a rectangle with sides $(55+39)\,.\,(55+56)$ which is subdivided into 13 sub-squares. *A dissection into squares in which a combination of the subset of the*

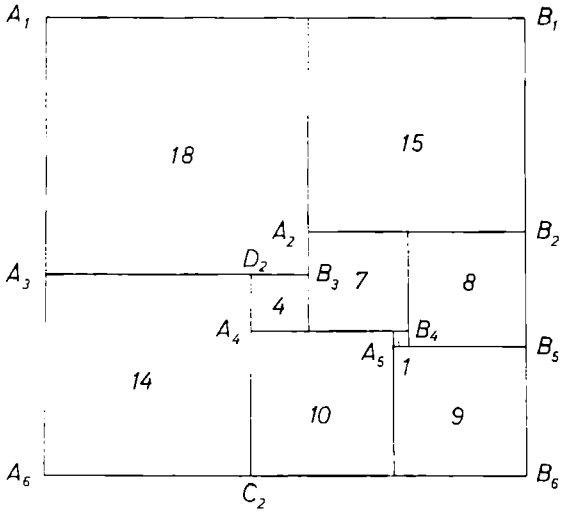


FIG. 30

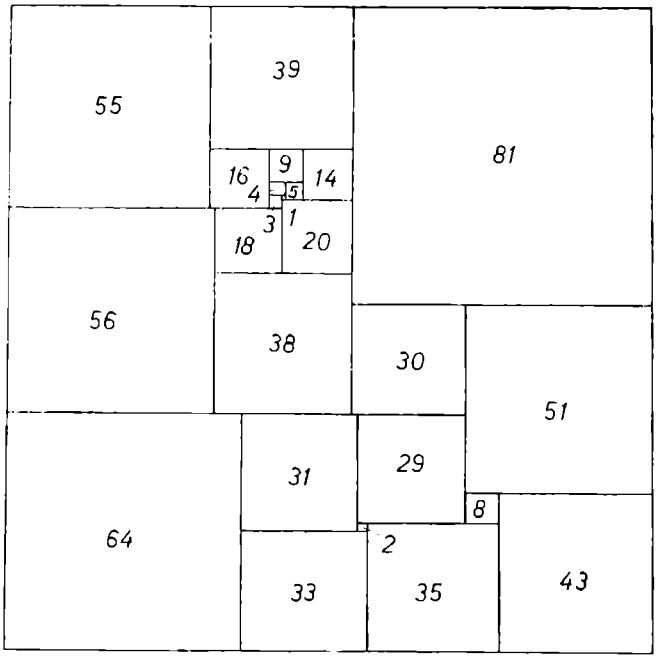


FIG. 31

squares to form a rectangle is impossible is called " simple ". The first decomposition into 55 sub-squares stated by R. SPRAGUE was not simple, nor was that into 26 sub-squares stated in [VII 7]. The smallest number n, for which a simple decomposition is known at present is 38. The coding of this decomposition is quoted in (3). Decompositions of rectangles into squares which are *all different* are called " perfect " decompositions. So far we have not taken any others into account. However, it makes perfectly good sense to ask about *simple, imperfect* decompositions. Fig. 32 shows a decomposition of this kind, which is *not perfect but simple*, of a square into 13 sub-squares.

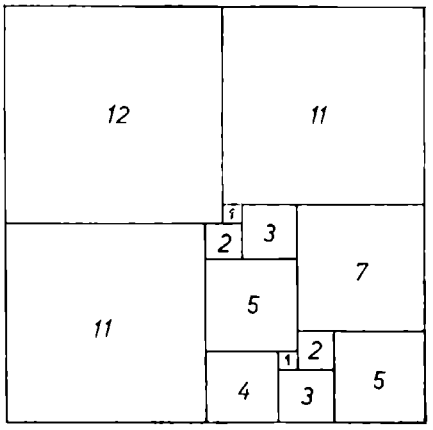


FIG. 32

The following questions now suggest themselves for our consideration:

(6) *Which is the smallest number n (n', n'') of squares used*

(6a) *in a perfect (simple, simple and perfect) decomposition of*

(6b) *a square?*

(6c) *Which is the smallest number n_R of sub-squares in a perfect decomposition of rectangles?*

(6d) *For which rectangles is a perfect decomposition into sub-squares possible?*

The number n (problem **(6)**) is still unknown and so is n''. On the other hand n' is known. This number is 13:—The decomposition which is shown in fig. 32 and which is simple but imperfect cannot be " improved ". The problems **(6d)** and **(6c)** are also solved. The answers to these are given at the end of this chapter.

2. The Definition of a Graph

A series of results from our range of problems was found obviously by untiring investigations. However, it is natural to look for a process by which one can discover all the possible decompositions of rectangles. As a first step in this direction one can try to limit the number of possibilities to be investigated.

For this purpose a graphical representation of the decomposition is introduced, this being suggested by the code-method described

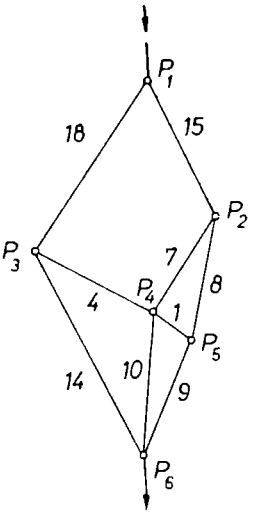


FIG. 33a

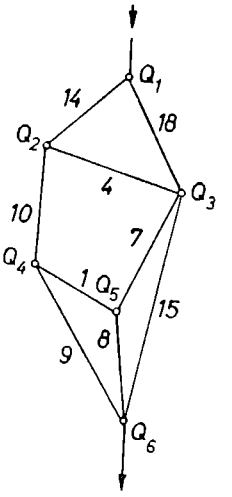


FIG. 33b

on p. 92. There the squares which were associated with the same horizontal cut, were collected in brackets. Now, if one replaces the horizontal cuts AB by points P_ν (e.g. the mid-point of each horizontal cut) and replaces each square by a line joining these points then one obtains a graph, e.g. the decomposition of fig. 30 is represented schematically in fig. 33a. The numbers attached to the lines are the lengths of the sides of the corresponding squares. The lengths of the lines of the graph are immaterial for our purposes. These depend only on the topological structure (and the noted lengths of the sides of the squares). It is certainly useful to represent the points of the graph so that (in a suitable orientation of the drawing) P_ν is placed " higher " than $P_{\nu+1}$.

For the later, systematic search for dissections of the rectangle it is significant that in a graph with V vertices only the first and last points (P_1 and P_V) can be of order 2. An interior point $P_m (1 < m < V)$

of order 2 would be possible only if two equal squares have a side in common. However, this is expressly excluded. Therefore we have for the vertices and edges of our graph the inequality

$$3V \leqq 2E+2. \tag{4}$$

Hence, from Euler's theorem for nets (Theorem 1, p. 9)

$$V-E+A = 1,$$

it follows for the number A of areas (" mesh units ")

$$A \geqq \frac{E+1}{3}. \tag{5}$$

Now one can introduce similarly the vertical cuts of a decomposition of a rectangle. These are the lines formed from the vertical sides of squares lying in a straight line. In fig. 30 only one vertical cut, C_2D_2, is indicated. Each vertical cut joins two points of two horizontal cuts $A_\rho B_\rho$ and $A_\sigma B_\sigma$ and the squares associated to the left and right of the vertical cut correspond to straight lines which join P_ρ and P_σ in the graph. Hence there is established a unique relation between the *interior* vertical cuts and the areas of the graph. The number A appearing in (4) and (5) is also the number of *interior* vertical cuts. If we take into account the two vertical sides of the given rectangle then we have altogether $v = A+2$ vertical cuts.

Now one can, of course, correlate to the rectangle, oriented according to our convention, yet a second graph in which the points correspond to the vertical cuts of the decomposition. The lines of this graph which is *dual* to the first graph again correspond to the sub-squares, but the mesh units of the net here correspond to the horizontal cuts. Fig. 33b shows the graph which is dual to fig. 33a.

If we denote the vertices, edges and areas of the second graph (dual to the first) by V_1, A_1, E_1 then we have, analogous to (4) and (5),

$$3V_1 \leqq 2E+2 \tag{4'}$$

$$A_1 \geqq \frac{E+1}{3}, \tag{5'}$$

where we have already allowed for the number of edges in both graphs being equal: $E = E_1$. Further the number of vertices V_1 in the second graph is equal to the number of vertical cuts, therefore

$$V_1 = A+2 = v. \tag{6}$$

Similarly we have for V

$$V = A_1 + 2 = h, \tag{7}$$

where h is the number of horizontal cuts in the given decomposition.

Finally, from (4), (4'), (5), (5'), (6) and (7) one obtains for the number of mesh units of both graphs the estimate

$$\frac{E+1}{3} \leqq \left\{ \begin{matrix} A \\ A_1 \end{matrix} \right\} \leqq \frac{2E-4}{3}. \tag{8}$$

This inequality together with Euler's equation (Theorem 1) makes it possible to determine the possible numbers for the areas (" mesh units ") and vertices of the graph for the decomposition of rectangles into a given number $E = E_1$ of squares. Since only integral values come into the question, we have collected for $E = 5, 6, \ldots, 13, 14$ the possibilities to be considered in the following table:—

E	A	V	E	A	V
5	2	4	10	4; 5	7; 6
6	—	—	11	4; 5; 6	8; 7; 6
7	3	5	12	5; 6	8; 7
8	3; 4	6; 5	13	5; 6; 7	9; 8; 7
9	4	6	14	5; 6; 7; 8	10; 9; 8; 7

3. Some Calculations for Decompositions

Now we must investigate which of the graphs possible from the table actually arise in a dissection of rectangles. In this connection we note the rules to which the side numbers of the squares are subject. These numbers are always written by the side of the lines of the graph (figs. 33a and 33b). The squares " sitting on " a horizontal cut give the same sum of side lengths as those " lying below " such a cut. The corresponding statement is true for the squares placed to the right and left of a vertical cut. Since the horizontal cuts correspond to the points and the (interior) vertical cuts to the " mesh units " of the graph, there thus arise the rules for the side lengths of the square.

In order to be able to formulate these rules conveniently we agree that a *negative* side number $x_{\rho\sigma}$ is correlated to a line $P_\rho P_\sigma$ of a graph if $\rho > \sigma$ and a positive number if $\rho < \sigma$. Therefore, in fig. 33a, we have written down the number 4 at the line P_3P_4; accordingly -4 would be the corresponding side number for P_4P_3.

Further we want to say that the side numbers which are correlated to all the graph lines $P_\rho P_{\sigma_\nu}$ intersecting at P_ρ belong to an interior point P_ρ of a graph. Similarly, we correlate a mesh unit $P_{n_1}P_{n_2}\dots P_{n_\tau}P_{n_1}$ of the graph with all the numbers $x_{n_\nu n_{\nu+1}}$ for the lines $P_{n_\nu}P_{n_{\nu+1}}$, always observing the rule for signs. Then there follows at once from the definition of side numbers and the rule of signs:

The algebraic sum of the side numbers corresponding to an interior point or to a mesh unit is equal to 0*:*

$$\sum_{\nu=1}^{g} x_{\rho\sigma_\nu} = 0, \quad \sum_{\nu=1}^{h} x_{n_\nu n_{\nu+1}} = 0, \quad n_{h+1} = n_1, \tag{9}$$

where g is the order of the point, h the number of vertices of the mesh unit.

Now the search for decompositions of rectangles (or possibly of squares) can continue as follows:—One draws the graphs which are topologically possible from the table on p. 97 and tests whether a " practicable " solution of the diophantine equations (9) exists for them. For example, solutions for which one $x_{ik} = 0$ are not " practicable ". Also solutions with $x_{ik} = x_{i\,k+1}$ are excluded because they would lead to a dissection with two equal, adjacent squares. Then this shows very quickly that most of the graphs which are topologically possible do not give practicable dissections.

Let us take as an example the first case of our table $E = 5$ (fig. 34). One can easily see directly that this figure cannot be the graph of dissection of a rectangle. However, we want to derive the reason from our statement about the algebraic sum of the side numbers in order to show how one can also achieve one's purpose in complicated cases.

Let it be assumed (according to our convention about the orientation of the rectangle) that $x_{13} > x_{12}$. Then we must have

$$x_{21}+x_{23}+x_{24} = 0, \quad x_{31}+x_{32}+x_{34} = 0, \tag{10}$$
$$x_{13}+x_{34}+x_{42}+x_{21} = 0.$$

Hence, it follows that

$$x_{12} = -x_{21} = x_{23}+x_{24} > x_{24}, \quad x_{13} = x_{34}-x_{23} < x_{34},$$

and, since $x_{13} > x_{12}$, therefore

$$x_{12}+x_{24} < 2x_{12} < 2x_{13} < x_{13}+x_{34}$$

in contradiction to the 3rd equation of (10).

Now one can investigate the other graphs in a similar way. The systematic investigation of the graphs for $E = 6, 7, \dots$ is not

particularly difficult but quite lengthy work. C. J. BOUWKAMP has carried out this investigation completely for $E = 5, 6, \ldots, 13$. There it turns out that practicable solutions are possible first for $E = 9$. Therefore there always appear for dissections of rectangles into squares of different side lengths at least 9 squares. Fig. 30 shows an example of such a dissection with $E = 9$.

With increasing E there are always more possibilities. In the work of BOUWKAMP the enumeration of the possible solutions by his code-procedure for $E = 13$ occupies several printed pages. There are ([VII 8], p. 1299) 585 decompositions of rectangles into 9, 10, 11, 12 or 13 incongruent squares. Among these rectangles there is *not a single square*. On the other hand there also exists among the 72 imperfect (but not trivial) dissections for $9 \leqq E \leqq 13$ the simplest "simple" (but not perfect) dissection into 13 squares which is represented in fig. 32.

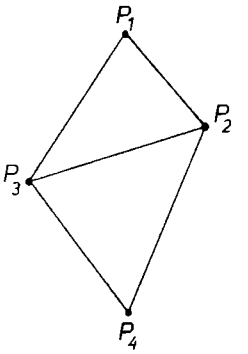


FIG. 34

One can build perfect dissections of squares from simpler dissections of rectangles. The method was already used by R. SPRAGUE in the construction of his first example [VII 4]. The basic idea is the following: One decomposes a square (fig. 35) into two unequal squares $\mathcal{Q}_1$ and $\mathcal{Q}_2$ and two rectangles $\mathcal{R}_1$ and $\mathcal{R}_2$. If one succeeds in decomposing the rectangles $\mathcal{R}_1$ and $\mathcal{R}_2$ into squares which are all different so that none of the squares of $\mathcal{R}_1$ is equal to a square of $\mathcal{R}_2$ and, in addition, all sub-squares of $\mathcal{R}_\nu$ ($\nu = 1, 2$) are different from $\mathcal{Q}_1$ and $\mathcal{Q}_2$, a perfect (but not simple) decomposition of the whole square is obtained. By means of two different decompositions of the rectangle with sides in the ratio 13:16 R. SPRAGUE obtained a decomposition of the square into 55 sub-squares in this way. Using the results of BOUWKAMP one can also obtain by this method a decomposition into 28 incongruent sub-squares, namely: there exist two different dissections of rectangles with sides in the ratio 593:422 into 13 different sub-squares:—

$$\begin{aligned} \mathcal{R}_1&: (222,164,207)\ (80,41,43)\ (39,2)\ (37,215)\ (200,22)\\ &\qquad\qquad (78) \\ \mathcal{R}_2&: (247,154,192)\ (116,38)\ (230)\ (175,72)\\ &\qquad (49,67)\ (103,18)\ (85). \end{aligned} \tag{11}$$

There are still some other methods (e.g. [VII 8], p. 74) of obtaining a perfect partitioning of the square from decompositions of rectangles. All decompositions found in this way are not simple. In the paper [VII 7] a method is stated for finding simple and perfect decompositions of squares. The method is certainly quite complicated and leaves open the question of which is the smallest number n'' of sub-squares which appear in a simple and perfect decomposition.

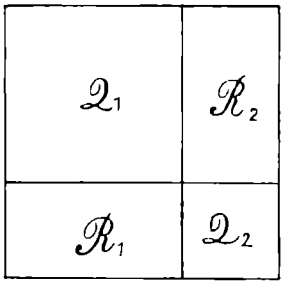


FIG. 35

4. THE INTERPRETATION IN ELECTRICAL ENGINEERING

It was already mentioned on p. 92 that the dissection of rectangles into incongruent squares is of significance for a problem in electrical engineering. We have kept the technical formulation separate from the mathematical considerations but we do not want to omit this interesting interpretation of our problem.

Let electrodes be placed at two opposite (horizontally oriented) sides of a homogeneous metal plate of rectangular form. Let a constant electrical current flow from one electrode to another. We consider the unit of current and of voltage to be chosen so that the horizontal sides of the rectangle are just equal to the current and the vertical equal to the voltage. Then the current lines run vertically and the potential lines horizontally. Therefore there is no change in the behaviour of the current if we construct vertical cuts in the plate and insert along a horizontal line in the plate a conductor of resistance zero.

Let us suppose that our rectangle is decomposed into squares which are all incongruent. Now we cut the plate at the vertical sides of the sub-squares and apply electrodes of infinitely high conductivity at the horizontal sides of the plate. Thus the current is not changed. The ratio $V:J$ (voltage:current) is, with our choice of units, just equal to 1 in *each* of the sub-squares. Therefore each sub-square represents a resistance of magnitude 1. Now one can concentrate all the horizontal electrodes at a point and roll together the sub-squares to form a wire. Then there results from the rectangular plate an electrical network which we have already represented schematically in the graph. Then each non-trivial

decomposition of a rectangle corresponds to a network of unit resistances in which no simple parallel or series circuits occur. If the rectangle to be decomposed is the particular case of a square, then one obtains in this way an electrical network formed of resistances of x ohms where the whole network also has x ohms resistance. The rules (9) can then be interpreted as Kirchoff's laws for electrical networks.

For the problem in electrical engineering it is unimportant whether the sub-squares appearing have all different side lengths or not. What matters is only that the decomposition is non-trivial when none of the squares has a side in *common*. Thus the graph belonging to the decomposition in fig. 32 represents the simplest net formed of resistances of magnitude 1 and itself having resistance 1.

5. The Decomposition of Rectangles with Commensurable Sides

Dehn showed [VII 1] already in the first work on decomposition of rectangles that dissection into squares which are all unequal is possible only for rectangles with commensurable sides. It was quite a long time before anyone succeeded in proving that each rectangle with commensurable sides is also in fact decomposable in the prescribed way [VII 5].

We restrict ourselves to sketching the basic ideas of a method of decomposition of this kind. Firstly, one shows that the unit square can be decomposed in infinitely many different ways into incongruent squares, where it turns out that in a series

$$\mathscr{Z}_1, \mathscr{Z}_2, \mathscr{Z}_3, \ldots \tag{12}$$

of decompositions of this kind each square in $\mathscr{Z}_\nu$ is different from each square in $\mathscr{Z}_\mu$. One can obtain a series (12) of decompositions e.g. by starting from a rectangle $\mathscr{R}$ which can be decomposed in two different ways into incongruent squares. Those rectangles assigned with the decomposition into squares we call $\mathscr{R}_1^0$ and $\mathscr{R}_2^0$.†

Now one obtains a series of pairs of rectangles $\mathscr{R}_1^\nu$ and $\mathscr{R}_2^\nu$ on applying to one of the smaller sides of $\mathscr{R}_1^{\nu-1}$ or $\mathscr{R}_2^{\nu-1}$ respectively a similar map of these rectangles (fig. 36).‡

† R. Sprague has used in his work [VII 5] a pair of rectangles which yields a square decomposed into 55 sub-squares. Using example (11) one gets a square which is decomposed into 28 incongruent sub-squares.

‡ Also other methods of constructing perfect squares from decomposed rectangles can be used to construct a series (12) e.g. those stated by Bouwkamp ([VII 8]) p. 74).

Besides, in each of the rectangles, the decomposition into sub-squares must be plotted. From the pairs of rectangles are now constructed squares $\mathfrak{Q}_\nu$ ($\nu = 1, 2, 3, \ldots$) according to the pattern of fig. 35 and hence one obtains by a suitable similarity transformation a series (12) of decompositions of the unit square. Now it must be proved that the sub-squares appearing in the different decompositions $\mathfrak{Z}_\nu$ are different from one another.†

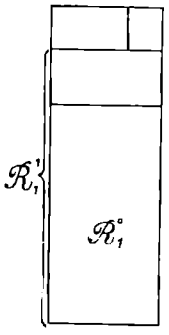


FIG. 36

In this way one has obtained the possibility of decomposing each rectangle with commensurable sides into squares which are all different. In fact:— One only requires to decompose the given rectangle by equidistant parallel lines into equal squares and in each of the squares q_ρ ($\rho = 1, 2, \ldots, m$) so obtained to apply the decomposition $\mathfrak{Z}_\rho$ ($\rho = 1, 2, 3, \ldots, m$). In this way one has obtained a perfect but, of course, not simple, decomposition of the given rectangle.

Of the problems (6), (6a), (6b), (6c) and (6d) formulated on p. 94 (6) and (6b) are still open; and the others are dealt with by the results of sections VII 3 and VII 5. For the numbers n, n' and n'' (p. 94) we have the statements:—

$$13 < n \leqq 24, \quad n' = 13, \quad 13 < n'' \leqq 38.$$

† We omit this proof. See [VII 5].

Chapter VIII

UNSOLVABLE EXTREMAL PROBLEMS

1. A Theorem on the Triangle

Every mathematician knows that there are infinite sets which have no greatest or no smallest element. If one poses, e.g., the problem of determining the smallest element of the set $\left\{\frac{1}{n}\right\}$ (n an integer greater than zero) then one has formulated an "unsolvable" problem. In this case it would be obvious indeed to say that one had posed a senseless question.

Now there exist extremal problems for which one would be inclined, on looking any of them over, to give the opinion that each must have a solution, but one finds on closer analysis of the problem that one is looking for the smallest (or greatest) element of a set $\mathcal{S}$ which, in fact, has no minimum (or maximum). The problem does not appear to nonsensical as that about the smallest number of the set $\left\{\frac{1}{n}\right\}$, and yet the statement of the problem is quite similar.

This is the case, e.g., for the problem (7) on p. 3. Here the set $\mathcal{S}$ is the set of surface areas of the domains in which the line of length 1 can be moved in the prescribed way. In fact, there exist domains of arbitrarily small (but different from zero) content, in which the rotation is possible. This must be proved in the following. In the proof of the theorem only the simplest facts of elementary geometry are used. Therefore the theorem could be made accessible to an interested fifth form boy.

First of all we prove a lemma on the triangle.

Lemma: *It is always possible to decompose a triangle ABC by transversals from C into sub-triangles CAA_1, CA_1A_2, ..., $CA_{n-1}B$ so that the union of the set (arising from a suitable parallel translation of the sub-triangles) of displaced sub-triangles has arbitrary small area.*

To prove this we divide the side $AB = c$ of the triangle ABC into 2^m equal parts (fig. 37a), where the number m can be chosen arbitrarily large. Let the dividing points lying between A and B be $A_1, A_2, \ldots, A_{2^m-1}$. Then we draw $m+1$ equidistant lines which are parallel to AB so that the height h_c of the triangle ABC is divided into $m+2$ parts. We number the parallel lines g_ν ($\nu = 0, 1, 2, \ldots, m$) so that the suffix increases with the distance from C.

Now we collect the 2^m sub-triangles into 2^{m-1} pairs:—The first pair is $\Delta_1 = AA_1C$ and $\Delta_2 = A_1A_2C$, the next $\Delta_3 = A_2A_3C$,

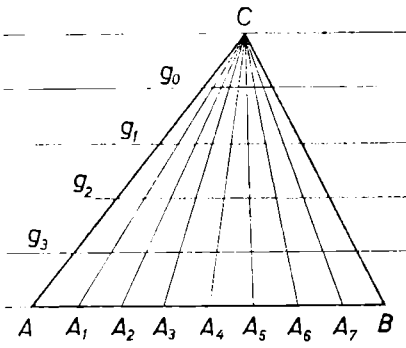


FIG. 37a

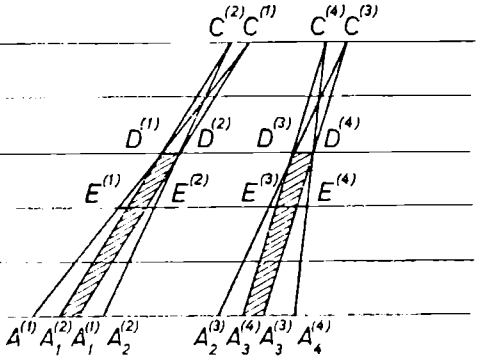


FIG. 37b

$\Delta_4 = A_3A_4C$, and so on. A parallel displacement of these pairs is carried out (fig. 37b) so that the displaced triangles cut the same segment on g_1. In fig. 37b only the displacements for the first two pairs are indicated. The triangles $A^{(1)}A_1^{(1)}C^{(1)}$ and $A_1^{(2)}A_2^{(2)}C^{(2)}$ have, for example, the line $D^{(1)}D^{(2)}$ as their common intersection with g_1. The surface area of the figure resulting from the translation is smaller than the sum of the surface areas of the two separate triangles because the two displaced triangles have, of course, a parallelogram in common. In fig. 37b the parallelograms

$$D^{(2)}D^{(1)}A_1^{(2)}A_1^{(1)} \text{ and } D^{(4)}D^{(3)}A_3^{(4)}A_3^{(3)}$$

belonging to the first two pairs are shaded in.

Now the intersection $d(g_1) = DD'$ of the triangle ABC with g_1 is, by the theorem on rays, equal to

$$d(g_1) = \frac{2c}{m+2},$$

and the intersection of the 2^{m-1} pairs with g_1 is therefore

$$d^{(1)}(g_1) = \frac{c}{m+2},$$

since, by the displacement, just two equal sub-lengths of the intersection are always made to coincide.

We now calculate or find a limit for $d^{(1)}(g)$ for the other parallel lines. For the parallels lying between C and g_1 obviously

$$d^{(1)}(g) \leqq d(g) < d(g_1) = \frac{2c}{m+2}. \tag{1}$$

From the translation of the triangles we have, in addition, for the parallels lying between g_1 and AB

$$d^{(1)}(g) = d(g) - d^{(1)}(g_1) = d(g) - \frac{c}{m+2}. \tag{2}$$

In particular, for g_2 we have

$$d^{(1)}(g_2) = d(g_2) - \frac{c}{m+2} = \frac{3c}{m+2} - \frac{c}{m+2} = \frac{2c}{m+2}, \tag{3}$$

and for the parallels lying between g_1 and g_2

$$d^{(1)}(g) < d(g_2) - d^{(1)}(g_1) = \frac{2c}{m+2}. \tag{1'}$$

Therefore (1) is true for *all* parallels between C and g_2.

Now the pairs of triangles are displaced into one another in groups of fours. This must be done so that each two displaced pairs cut the same length on the parallel line g_2. The first two pairs indicated in fig. 37b are, for example, to be displaced so that the lines $E^{(1)}E^{(2)}$ and $E^{(3)}E^{(4)}$ are made to coincide. Since the other translations occur in the same way, the surface area of the 2^{m-2} groups of four is smaller than that of the 2^{m-1} pairs of triangles and the intersection $d^{(2)}(g)$ of the line parallel to AB with the triangles now collected into groups of four is likewise reduced by this displacement process

$$d^{(2)}(g) \leqq d^{(1)}(g) \leqq d(g). \tag{4}$$

In particular, the intercept of g_2 with the groups of four becomes

$$d^{(2)}(g_2) = \tfrac{1}{2} d^{(1)}(g_2) = \frac{c}{m+2}. \tag{5}$$

The parts of the pairs of triangles lying between g_2 and AB are combined by the second displacement into trapeziums in which a parallelogram is twice covered (fig. 38). Thus e.g. $A^{(1)}A_2^{(2)}E^{(2)}E^{(1)}$ and $A_2^{(3)}A_4^{(4)}E^{(4)}E^{(3)}$ (fig. 37b) are displaced together into the

quadrangle $A^{(5)}A_4^{(6)}E^{(6)}E^{(5)}$ (fig. 38), where $E^{(2)}A_2^{(2)}$ goes into $E^{(6)}A_2^{(6)}$ and $E^{(3)}A_2^{(3)}$ into $E^{(5)}A_2^{(5)}$. The parallelogram $A_2^{(5)}A_2^{(6)}E^{(6)}E^{(5)}$ is therefore covered twice. This means that by (5) and (2) for parallels g lying between g_2 and AB

$$d^{(2)}(g) = d^{(1)}(g) - d^{(2)}(g_2) = d^{(1)}(g) - \frac{c}{m+2} = d(g) - \frac{2c}{m+2}.$$

For the straight lines g lying between g_2 and g_3, in particular, it is true that

$$d^{(2)}(g) < d(g_3) - \frac{2c}{m+2} = \frac{4c}{m+2} - \frac{2c}{m+2} = \frac{2c}{m+2}. \qquad (6)$$

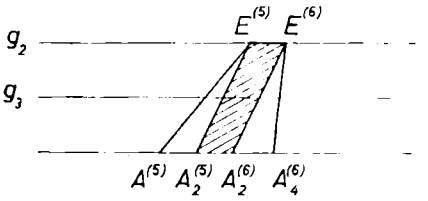


FIG. 38

From (1′), this inequality (6) is true for all parallels between C and g_3.

The displacement process is now continued. At the μth step $2^{m-\mu+1}$ domains (which each consist of $2^{\mu-1}$ displaced triangles) are translated so that a pair of these domains always has the same interval in common with the straight line g_μ. At the last stage one obtains finally a last domain $\mathfrak{D}_m$ from two domains. From the argument used above one deduces, in addition, that

$$d^{(\mu)}(g) < \frac{2c}{m+2}$$

is true for all parallel lines between C and $g_{\mu+1}$. Then in $\mathfrak{D}_m$ we have

$$d^{(m)}(g) < \frac{2c}{m+2} \qquad (7)$$

for *all* the parallel lines.†

Hence we have a simple, practicable way to estimate the surface area of the last domain $\mathfrak{D}_m$. Of course, $\mathfrak{D}_m$ is decomposed by the

† The complete proof of (7) can be produced easily by mathematical induction using the given method. The details are left to the reader.

parallel lines into trapeziums and triangles. However, the area of the trapeziums and triangles can be calculated from the product of the median and height.

For the area $a(\mathscr{D}_m^{(\delta)})$ of the part of $\mathscr{D}_m$ lying between any two parallel lines g_μ and $g_{\mu+1}$ (at distance δ apart), from (7), it is true that

$$a(\mathscr{D}_m^{(\delta)}) < \frac{2c\delta}{m+2}.$$

Hence we have for the area of $\mathscr{D}_m$ itself

$$a(\mathscr{D}_m) < \frac{2ch_c}{m+2}.$$

By choosing a large number m we can therefore deduce that $a(\mathscr{D}_m)$ becomes arbitrarily small. And this is what we had to prove.

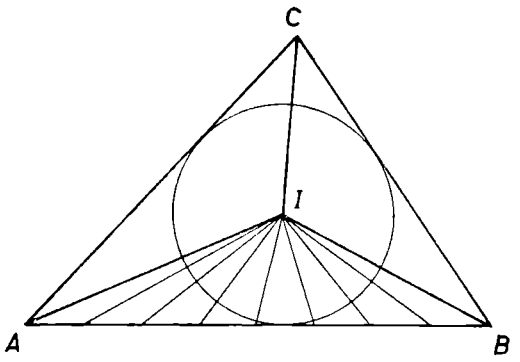


FIG. 39

2. BESICOVITCH'S THEOREM

Using the lemma just proved we can demonstrate the unsolvability of the extremal problem (7) (p. 3); namely that the theorem, first proved by BESICOVITCH, is true:—

Theorem 27: *There are domains of arbitrarily small surface area in which a line UV of length* 1 *can be moved by rotations and translations so that the direction angle between UV and a fixed line is turned through* 360°.

To prove this we use any triangle ABC with incircle of radius 1. Let I be the incentre (fig. 39). Each of the sub-triangles ABI, BCI and CAI are now divided into 2^m sub-triangles with vertex I in the way described in the proof of the lemma in Section VII 1. One chooses m large enough to ensure that each of the domains $\mathscr{D}_m^{(1)}$, $\mathscr{D}_m^{(2)}$, $\mathscr{D}_m^{(3)}$ corresponding to the triangles ABI, BCI and CAI has

an area which is smaller than $\varepsilon/6$, where ε is an arbitrarily small, given, real number. We can combine these three domains to form a domain $\mathscr{D}_m$, say, by the following rule:—$\mathscr{D}_m^{(2)}$ is adjoined to $\mathscr{D}_m^{(1)}$ by joining the first sub-triangle of BCI to the last sub-triangle of ABI. Similarly, $\mathscr{D}_m^{(3)}$ is adjoined to $\mathscr{D}_m^{(2)}$. Thus we obtain a domain $\mathscr{D}_m$ whose area is at most equal to $3 \cdot \varepsilon/6 = \varepsilon/2$.

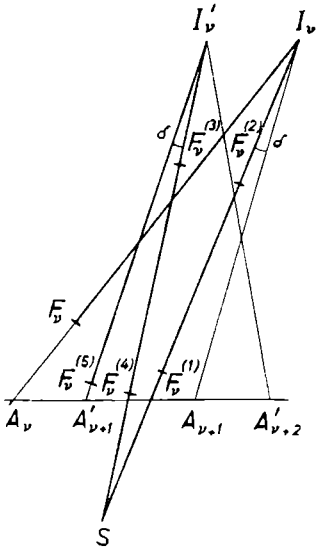


FIG. 40

Now the basic idea of the proof is as follows:—One can move a line of length 1 through 360° around the incentre I in the triangle ABC. Now this motion must not be carried out in the triangle ABC itself but in the domain $\mathscr{D}_m$ arising from the translations and adjoining of domains. As we shall show this can be achieved only if, using parallel displacements of the line, we leave this domain. It will then be shown, however, that there is a domain $\mathscr{D}_m^*$ (containing $\mathscr{D}_m$) of surface area

$$\varepsilon_1 \leqq 2 \cdot \frac{\varepsilon}{2} = \varepsilon$$

in which rotations *and* translations can be performed.

Let $A_\nu A_{\nu+1} I_\nu$ and $A'_{\nu+1} A'_{\nu+2} I'_\nu$, say, (fig. 40) be two triangles from $\mathscr{D}_m$ which are obtained by displacing neighbouring triangles of ABI (fig. 39), BCI or CAI. Let $I_\nu F_\nu$ be the line of length 1 to be moved and $I_\nu S$ a ray, making with $A_{\nu+1} I_\nu$ an angle δ which can be chosen arbitrarily small. Now let this angle be formed at I'_ν on $I'_\nu A'_{\nu+1}$ and, of course, in the half plane in which I_ν lies. Since $I'_\nu A'_{\nu+1}$ and $I_\nu A_{\nu+1}$ are parallel, the free arms of the two angles δ will intersect at S, say. We now move our line of length 1 from the position $I_\nu F_\nu$ on the side $A_\nu I_\nu$ of the triangle $A_\nu A_{\nu+1} I_\nu$ to the corresponding position $I'_\nu F_\nu^{(5)}$ on the side $I'_\nu A'_{\nu+1}$ of the triangle $A'_{\nu+1} A'_{\nu+2} I'_\nu$. The motion is carried out in several steps using rotations and translations (see fig. 40):—

1. $I_\nu F_\nu \to I_\nu F_\nu^{(1)}$ (Rotation about I_ν)

2. $I_\nu F_\nu^{(1)} \to F_\nu^{(2)} S$ (Translation)

3. $F_\nu^{(2)} S \to F_\nu^{(3)} S$ (Rotation about S)

4. $F_\nu^{(3)}S \to I'_\nu F_\nu^{(4)}$ (Translation)

5. $I'_\nu F_\nu^{(4)} \to I'_\nu F_\nu^{(5)}$ (Rotation about I'_ν).

This process can be continued through all the sub-triangles of $\mathfrak{D}_m$. In this way the line, in fact, experiences a rotation through 360°. There only remains to be considered whether the path of the line may also be imbedded in a domain of arbitrarily small surface area. $\mathfrak{D}_m$ has, of course, an area which is smaller than $\varepsilon/2$. However, the line is moved (partially at least) *outside* of this domain

1. in the displacement $I_\nu F_\nu^{(1)} \to F_\nu^{(2)}S$ or $F_\nu^{(3)}S \to I'_\nu F_\nu^{(4)}$,

2. in the rotation through the angle 2δ about S.

In this rotation about S the line sweeps out a sector of area δ. The area of all the $3 \cdot 2^m$ sectors is therefore $3 \cdot 2^m \cdot \delta$. If we choose

$$\delta < \frac{\varepsilon}{2^{m+2} \cdot 3},$$

then this area becomes smaller than $\varepsilon/4$. There still remain the lines $F_\nu^{(1)}F_\nu^{(2)}$ or $F_\nu^{(3)}F_\nu^{(4)}$ which (in fig. 40 this is not the case) can lie partially outside $\mathfrak{D}_m$. Let us suppose that l is the total length of all such lines. Then we can imbed these lines in rectangular strips of arbitrarily small breadth η. If we choose $\eta < \frac{\varepsilon}{4l}$ then

$$l \cdot \eta < \frac{\varepsilon}{4}.$$

Hence it is clear that the content of the union $\mathfrak{D}_m^*$ of sets $\mathfrak{D}_m$, the parallel strips and the sectors becomes arbitrarily small:—

$$a(\mathfrak{D}_m^*) \leqq a(\mathfrak{D}_m) + l\eta + 3 \cdot 2^m \cdot \delta < \frac{\varepsilon}{2} + 2\frac{\varepsilon}{4} = \varepsilon.$$

3. Sets of Points at an Integral Distance Apart

Recently in a mathematical journal [VIII 1] the following question was posed:—

(7a) *For the Euclidean space of dimension d what is the maximal number m_d of points which are at an integral distance apart from each other and do not all lie on the same straight line?*

To elucidate this problem we must make use of a theorem of CHOQUET and KREWERAS:—

Theorem 28: *If an infinite set of points $\mathcal{S}$ of the plane has the property that the distance between any two points of the set is integral, all the points of $\mathcal{S}$ lie on a straight line l.*

To prove this theorem we denote by l the straight line determined by any two points A and B of the set $\mathcal{S}$. Now let C be any point of $\mathcal{S}$ which does not lie on l and let $AC \geqq BC$, say. Then the difference of the distances AC and BC

$$d_1 = AC - BC \; (< AB = d)$$

must also be integral. Thus the only possibilities for d_1 are the numbers 0, 1, 2, ..., $d-1$. Therefore C (and every other point $D \in \mathcal{S}$ but not lying on l) must lie on one of the d hyperbolae which are determined by the foci A and B and the transverse axis $a = \frac{1}{2}\delta$ ($\delta = 0, 1, 2, \ldots, d-1$).†

However, a corresponding argument is also possible for the line segment $AC = d^*$ determining the straight line l^*. Each of the points of $\mathcal{S}$ different from A and C must lie either on l^* or on one of the hyperbolae which are determined by the foci A and C and the axis $a^* = \frac{1}{2}\delta^*$ ($\delta^* = 0, 1, 2, \ldots, d^*-1$). But there are only finitely many points of intersection of the two straight lines l, l^* and the $d+d^*$ hyperbolae. However, since $\mathcal{S}$ has to be an infinite set, the supposition that a point C does not lie on the straight line l is false.

It is not difficult to extend this theorem, which is valid for the plane, to Euclidean spaces of higher dimension. Therefore the question noted above as number (**7a**) is a natural one and yet this problem is also an unsolvable extremal problem. It will be shown that, even in the plane, there is *no maximum* m_2 for the number of points which have integral distances between each two (without all of them lying on a straight line l). In other words:—There are in the plane an arbitrarily large number of points $P_1, P_2, \ldots, P_n$ for which all the distances $P_\nu P_\mu$ are integral without all the P_ν lying on a straight line. First of all Theorem 28 holds for an *infinite* set. We prove in particular

Theorem 29: *There exists in the plane an arbitrarily large number of points $P_1, P_2, \ldots, P_n$ lying on a circle for which the distance $P_\nu P_\mu$ between any two points is integral.*

† For $\delta = 0$ the hyperbola degenerates into the perpendicular bisector of AB.

To prove this we start with any isosceles triangle with sides of integral length. Of course, one can obtain any number of triangles of this kind by adding a Pythagorean triangle (a right-angled triangle with sides of integral length) to its reflection in a side which is not a hypotenuse. Starting from any point P_1 on the circumcircle $\mathscr{C}$ of a triangle ABC of this kind we now measure off as a chord one of the equal sides of ABC arbitrarily often. This produces points which are all different. One can prove this for example using the lemma proved on p. 135. Then any two of the points $P_1, P_2, \ldots, P_n$ obtained in this way have a *rational* distance apart. We prove this by mathematical induction. First of all our statement is correct for the first three points since the triangle $P_1P_2P_3$ is congruent to the triangle ABC and therefore has sides of integral length. Now let us suppose that our statement is true for the points $P_1, P_2, \ldots, P_n$. Then we have to show that P_{n+1} is also at a rational distance from all the points $P_1, P_2, \ldots, P_n$. Since the sequence of points $P_1, P_2, \ldots, P_n$ is congruent to the sequence $P_2, P_3, \ldots, P_{n+1}$, we have only to prove that P_1P_{n+1} is also rational. However, this follows from Ptolemy's theorem † according to which the product of the diagonals in an inscribed quadrangle $P_1P_2P_3P_{n+1}$ is equal to the sum of the products of each of the two opposite sides:—

$$P_1P_{n+1} \cdot P_2P_3 + P_3P_{n+1} \cdot P_1P_2 = P_1P_3 \cdot P_2P_{n+1}.$$

From the induction hypothesis it therefore follows that P_1P_{n+1} is also rational.‡

Now it is not difficult to obtain a set of points which are at integral distances apart. Let Q be the least common multiple appearing in any of the rational distances

$$P_\nu P_\mu = \frac{p_{\nu\mu}}{q_{\nu\mu}}$$

and $\mathscr{C}(Q)$ the circle with centre U and radius rQ (where U is the centre and r the radius of $\mathscr{C}$); let $Q_1, Q_2, \ldots, Q_n$ be the points of intersection of the radii UP_ν with $\mathscr{C}(Q)$. Then $\{Q_1, Q_2, \ldots, Q_n\}$ is a set of points with all the distances $Q_\nu Q_\mu$ integral.

One easily sees that the radius rQ of the circle $\mathscr{C}(Q)$ increases without bound with the number n. At most two of the distances

† See e.g. [IV 2] p. 104.

‡ Another form of this proof by HADWIGER [IV 4] avoids using Ptolemy's theorem. MÜLLER [VIII 3] originally gave the above proof.

P_1P_2, P_1P_3, ..., P_1P_n can be equal. Therefore among the rational numbers $P_\nu P_\mu$ appear, with increasing n, some with arbitrarily large denominator. Therefore the radius rQ of the circle $\mathscr{C}(Q)$ increases without bound with the number n. This agrees with the statement of Theorem 28, according to which the points lie on a straight line if the set $\mathscr{S}$ contains infinitely many elements. F. STEIGER [VIII 4] has shown that one can always find for each dimension d of a Euclidean space, for $N>d$, a system of N points which are integral distances apart and which do not all belong to the same linear subspace of dimension $d-1$.

A problem which is still unsolved [VIII 5] and which should be mentioned in this connection is the following:—

(7b) *Is there a denumerable, infinite set which is everywhere dense in the whole plane and for which every two points have a rational distance apart?*

One can easily show that the denumerable set $\{P_n\}$ on the circle $\mathscr{C}$ (of radius r) used in the proof of Theorem 29 is everywhere dense. On the straight line the points with rational coordinates form such a set.

Another example of a denumerable plane set of points which are at rational distances apart is the set $\mathscr{S}_1$, to which the points of the (x, y)-plane

$$(0, +1), (0, -1), \left(\frac{p^2-q^2}{2pq}, 0\right),$$

belong, where p and q are any integers which are different from zero. The numbers of this set † which are denumerable on the x-axis are everywhere dense there and the distance between any two points of the set is rational. It has not yet been determined whether or not there is a set of this kind which is dense *in the whole plane*.‡

† R. SPRAGUE is the originator of this example.
‡ On this range of problems see also [VIII 6].

Chapter IX

CONSTRUCTIONS WITH RULER AND COMPASSES

1. Some Comments on the Use of Ruler and Compasses

In antiquity mathematicians were already concerned with constructions using ruler and compasses. There were construction problems of this kind which can be formulated quite simply but whose solutions were not found at that time. One of these is the trisection of the angle (see problem (8), p. 3) and another the " Delian problem ":—The Delians, so the tale goes, had turned to Apollo to request help against a plague and the deity had demanded that they must " double " a sacred altar to him. When a cube with edges of twice the length did not produce successful results, it became clear to them that by doubling the edges the volume of the cube had not been doubled but multiplied by eight. And because the Delians were unsuccessful in their efforts to construct a cube of twice the volume, they sent a delegation to Plato's Academy requesting the solution of the problem.

But the men of the Academy could not solve the problem and, indeed, in the subsequent centuries right up to the present time professionals and amateurs (in recent years only amateurs!) have tried repeatedly to solve the Delian problem or the related problem of the trisection of the angle. Today we know that both these classical problems are unsolvable if one is allowed to use only ruler and compasses in the construction as was customary in classical times.

To avoid misunderstandings it is certainly necessary to say exactly what may and may not be done with ruler and compasses. If today mathematical outsiders still produce the solution of one or another unsolvable construction problem, it is generally due to the fact that the rules for constructions " using ruler and compasses "

are not clearly understood. Therefore, we want to make quite clear exactly what is meant by this.

Let there be given n ($n \geqq 2$) points A_ν ($\nu = 1, 2, \ldots, n$) in a plane $\mathscr{P}$. The ruler may be used to draw lines through any two points A_ν and A_μ and the compasses to draw circles with centre at any point A_ν and radius $A_\lambda A_\mu$. Let the points of intersection of these lines and circles be $A_{n+1}, A_{n+2}, \ldots, A_{n+m}$, say. Then one can continue this process (a finite number of times!), taking now as given the points $A_1, A_2, \ldots, A_n, \ldots, A_{n+m}$ and constructing additional points which are intersections of lines and circles (or lines and lines or circles and circles).

It will be useful to add what may *not* be done with ruler and compasses. The *ruler* may not be provided with markings which one "inserts" between lines or between a line and a circle or between circles. One can, of course, make the convention that the use of an "insertion ruler" is to be allowed. However, one must be clear that this means introducing a new instrument of construction. If one refers simply to a ruler, one always means that the usual ruler without markings is being used. We shall be concerned exclusively with constructions in which the ruler is to be used in the way described here.†

The *compasses* may not be used to discover by trial and error which points of a given circle and a line are equally far apart. It may be that able use of the method of trial and error can yield the same accuracy or better than unskilful use "according to regulations" of the apparatus allowed. To investigate the possibilities in construction this is just as unimportant as the fact that a "genuine" straight line cannot be drawn by any means with a pencil.

On elucidating what may be done using ruler and compasses we did not exclude the possibility of drawing a circle with centre at any point and *arbitrary* radius. In fact, methods involving this play a role in the bisection and measuring off of an angle and in other elementary problems. However, we deliberately refrain from using this possibility throughout, since we can always (for at least two given points) find among the given or constructed points two whose joining line is suitable for use as the radius in a problem which is usually solved using an arbitrary radius.

† On the theory of the insertion ruler and other means of construction such as set square, parallel ruler and so on, see e.g. [IX 7].

When at least two points are given, it is also possible to designate a given or constructed line as " unit line " and to introduce a coordinate system. In this way one can translate all the steps of the construction into algebraic ones by the methods of analytical geometry. If one produces two lines, say those determined by the points A_ν and A_μ and A_ρ and A_σ, to intersect then one obtains the coordinates of the points of intersection A_ν by solving the system of the two linear equations of the lines $A_\nu A_\mu$ and $A_\rho A_\sigma$. Quadratic equations have to be solved to calculate the points of intersection of a circle and a line or of two circles. Therefore the coordinates of each point A_{n+1}, obtained from the points $A_1, A_2, \ldots, A_n$ by an admissible construction step, can be determined from the coordinates of the points A_ν $(\nu = 1, 2, 3, \ldots, n)$ by the four basic arithmetical operations and the process of extracting square roots.

At this stage it is useful to introduce the idea, common in modern algebra, of the " number field ". By this one means a set of numbers which may be determined from a given set of numbers by addition, subtraction, multiplication and division (by numbers other than zero). The rational numbers form a field and so do all the numbers of the form $a+b\sqrt{c}$, where a, b and c are rational. In this case one says that the field $\mathfrak{F}_1$ of numbers $a+b\sqrt{c}$ arises from the field $\mathfrak{F}_0$ of rational numbers by " adjoining " $\sqrt{c}$.†

We want to prove here that the numbers of the form $a+b\sqrt{c}$ form a field. To prove that addition, subtraction and multiplication of numbers of this form lead to numbers of the form $r_1+r_2\sqrt{c}$ is trivial. However, a quotient

$$q = \frac{a+b\sqrt{c}}{a_1+b_1\sqrt{c}}$$

may also be written in this way: $q = r_3+r_4\sqrt{c}$. One sees this immediately if one multiplies numerator and denominator of q by $a_1-b_1\sqrt{c}$. Then one obtains

$$q = \frac{(a+b\sqrt{c})(a_1-b_1\sqrt{c})}{a_1^2-b_1^2c} = \frac{aa_1-b_1bc}{a_1^2-b_1^2c} + \frac{ba_1-ab_1}{a_1^2-b_1^2c}\cdot\sqrt{c}.$$

If the coordinates of all the given points $A_1, A_2, \ldots, A_n$ belong

† General statements about adjunction are to be found in textbooks of algebra, e.g. [IX 1].

to the field $\mathfrak{F}_0$ of rational numbers then the coordinates of all points constructed from them, using ruler and compasses, belong to a field $\mathfrak{F}_m$ which is defined as follows:—$\mathfrak{F}_m$ is obtained from $\mathfrak{F}_{m-1}$ by adjoining a number $\sqrt{c_{m-1}}$, where c_{m-1} belongs to a field $\mathfrak{F}_{m-1}$. Finally, $\mathfrak{F}_1$ is obtained from $\mathfrak{F}_0$ by adjoining $\sqrt{c_0}$, where c_0 is rational.

A simple example will illustrate this. Let us start with 2 given points A_1 and A_2 whose coordinates we take to be (0,0) and (0,1). Now A_1A_2 is divided in the well-known way " in golden section ". The point of division A_3 has the coordinates $(0, -\frac{1}{2}+\frac{1}{2}\sqrt{5})$. Then these numbers 0 and $-\frac{1}{2}+\frac{1}{2}\sqrt{5}$ belong to the field $\mathfrak{F}_1$ which results from the field $\mathfrak{F}_0$ of rational numbers by adjoining $\sqrt{5}$.

Now one draws the circles with centres A_1 and A_3 and radius $A_1A_2 = 1$. Let us call the two points of intersection A_4 and A_5. Their coordinates

$$\tfrac{1}{2}(-\tfrac{1}{2}+\tfrac{1}{2}\sqrt{5}),\ \pm\sqrt{\tfrac{5}{8}+\tfrac{1}{8}\sqrt{5}}$$

belong to the field $\mathfrak{F}_2$ which results from $\mathfrak{F}_1$ by adjoining $\sqrt{c_1} = \sqrt{\frac{5}{8}+\frac{1}{8}\sqrt{5}}$, $c_1 \in \mathfrak{F}_1$.

In general, the coordinates of the given points $A_1, A_2, \ldots, A_n$ will not all belong to $\mathfrak{F}_0$. Then we start from the field $\mathfrak{F}$ which results from $\mathfrak{F}_0$ by adjoining the irrational coordinates of the points A_ν ($\nu = 1, 2, \ldots, n$). Then each construction using ruler and compasses starting from the points $A_1, A_2, \ldots, A_n$ leads to points $A_{n+1}, A_{n+2}, \ldots, A_{n+m}$ whose coordinates belong to a field $\mathfrak{F}_m$ which results from $\mathfrak{F}_{m-1}, \mathfrak{F}_{m-2}, \ldots, \mathfrak{F}_1, \mathfrak{F}$ by repeated adjunction of square roots.

Conversely, one can construct each expression of the form $a+b\sqrt{c}$ from given lines of length a, b and c using ruler and compasses. One learns this already in school.

We sum up in

Theorem 30: *Let $A_1, A_2, \ldots, A_n$ be n given points in the plane and $\mathfrak{F}$ the number field which results from the field $\mathfrak{F}_0$ of rational numbers by adjoining the coordinates of $A_1, A_2, \ldots, A_n$ in a certain coordinate system. Then using ruler and compasses one can obtain those points of the plane whose coordinates belong to a field $\mathfrak{F}_m$ which has the following properties:—$\mathfrak{F}_m$ results from $\mathfrak{F}_{m-1}$ by adjoining $\sqrt{c_{m-1}}$, where $c_{m-1} \in \mathfrak{F}_{m-1}$ and, finally, $\mathfrak{F}_1$ is obtained from $\mathfrak{F}$ by adjoining a number $\sqrt{c}$, where $c \in \mathfrak{F}$.*

2. Some Construction Problems

Now, how is one to decide in problem (**8**) (p. 3) or in the classical problems mentioned in the preceding section whether the condition of Theorem 30 is satisfied or not? First of all let us formulate the problem of the quinquesection (or trisection) of the angle somewhat differently.

Let $\angle\alpha = \angle ABC$ be the angle to be dissected and A_1 the intersection of the circle with centre B and radius BC and of the arm BA; in addition, let D be the foot of the perpendicular from A_1 onto BC. If we now choose BC as unit of length, we have $BD = \cos\alpha$. Our dissection problem will be solved if we succeed in constructing a point E on the line BC with the property that $BE = \cos\alpha/5$. Conversely, if the quinquesection is obtained in any way by a ray BF then one can always construct $BE = \cos\alpha/5$. Therefore our problem is solved if and only if the point E can be constructed.

Now, we denote the unknown line segment $\cos\alpha/5$ by x and use the well-known formula †

$$\cos 5\beta = \cos^5\beta - \binom{5}{2}\cos^3\beta\sin^2\beta + \binom{5}{4}\cos\beta\sin^4\beta.$$

If one replaces $\sin^2\beta$ by $1-\cos^2\beta$ in this, it follows that

$$2\cos 5\beta = (2\cos\beta)^5 - 5(2\cos\beta)^3 + 5(2\cos\beta). \tag{1}$$

Hence, for $y = 2x$ and $5\beta = \alpha$, this becomes

$$y^5 - 5y^3 + 5y - 2\cos\alpha = 0. \tag{2}$$

If we denote by $\mathfrak{F}$ the field of numbers which results from the field $\mathfrak{F}_0$ of rational numbers by adjoining $\cos\alpha$, we can now formulate our problem as follows: *Is there a solution of* (2) *belonging to a field* $\mathfrak{F}_m$ *which is an extension of* $\mathfrak{F}$.‡

Before we turn to answering this question, we will establish the equations corresponding to (2) for some simpler problems and then advance by the easier problems to the answer to problem (**8**) (p. 3).

† See in this connection VI (7′).

‡ In this chapter by a field $\mathfrak{F}_m$ which is an extension of a field $\mathfrak{F}$ we always mean a field which is obtained from $\mathfrak{F}$ by successive adjunction of square root expressions as described in Theorem 30.

Let us denote by **(8a)** *the problem of the trisection of the angle.* Using the formula (see VI (7′))

$$\cos 3\beta = 4\cos^3\beta - 3\cos\beta$$

for the classical problem of the trisection of the angle, substituting $3\beta = \alpha$ and $y = 2x = 2\cos\beta$, one obtains the equation of third degree

$$y^3 - 3y - 2\cos\alpha = 0. \tag{3}$$

One is also led to an equation of third degree for the Delian problem of doubling the cube (problem **(8b)**). If the length of the side of the given cube is denoted by 1 and the unknown length of the cube with twice the volume by x, we have

$$x^3 - 2 = 0. \tag{4}$$

Finally, we want to deal with two construction problems which lead to equations of 4th degree:—

(8c) *It is required to draw (using ruler and compasses) an isosceles triangle with the radius of its circumcircle r and the radius of its incircle ρ.*

(8d) *It is required to construct (using ruler and compasses) an isosceles triangle with the radius of its circumcircle r and area J.*

In problem **(8d)** it is assumed that the area J is given by a line segment of length J.

For problems **(8c)** and **(8d)** we introduce the following notation: Let us call the base of the triangle to be constructed $AB = a$, the side b, the base angle β and the angle at the vertex α. We note that the circumcircle centre M lying on the axis of symmetry is characterised by the fact that the vertex angle between AM and the axis of symmetry is equal to α. Then we have

$$\sin\alpha = \sin(\pi - 2\beta) = \sin 2\beta = \frac{a}{2r} \tag{5}$$

and

$$\tan\frac{\beta}{2} = \frac{2\rho}{a}. \tag{6}$$

We put $\cos\beta = y$, $\frac{\rho}{r} = u$ and note that one can construct the distance u from ρ and r. Since

$$\tan\frac{\beta}{2} = \sqrt{\frac{1-\cos\beta}{1+\cos\beta}}$$

it follows from (5) and (6) that

$$y^2(1-y)^2-\frac{u^2}{4}=0.$$

This equation can also be written as follows

$$\left[y(1-y)+\frac{u}{2}\right]\left[y(1-y)-\frac{u}{2}\right]=0.$$

From this form one sees immediately that the equation can be solved in terms of square root expressions. Hence problem **(8c)** is essentially solved.

To examine problem **(8d)** we substitute

$$x=\sin\alpha=\frac{a}{2r}.$$

Then for the height h of the required triangle it is true that

$$h=r+r\cos\alpha.$$

Therefore the area becomes

$$J=\frac{a}{2}r(1+\cos\alpha)=r^2\sin\alpha(1+\cos\alpha). \tag{7}$$

If we write μ for the quotient $\frac{J}{r^2}$, we obtain from (7)

$$x^4-2\mu x+\mu^2=0. \tag{8}$$

We have now to examine whether the equations (2), (3), (4) and (8) can be solved in terms of square roots or not. For the equations of third degree one can answer this question fairly simply by using a theorem which we now consider.

3. The Proof of the Impossibility of (8a) and (8b)

Theorem 31: *If an equation of third degree with coefficients from a field $\mathfrak{F}$ has a solution which belongs † to a field $\mathfrak{F}_m$ which is an extension of $\mathfrak{F}$ then it has also a solution which belongs to $\mathfrak{F}$ itself.*

Let us suppose for the proof of this theorem that the equation

$$x^3+a_2x^2+a_1x+a_0=0 \tag{9}$$

† See in this connection footnote ‡ p. 117.

has coefficients from a field $\mathscr{F}$ and that a solution x_1 of this equation (9) belongs to a field $\mathscr{F}_m$ but not to the field $\mathscr{F}_{m-1}$. Then x_1 can be written as follows

$$x_1 = a+b\sqrt{c}$$

where the numbers a, b and c belong to the field $\mathscr{F}_{m-1}$. In the following we denote the number

$$x^* = A-B\sqrt{C} \quad (A, B, C \in \mathscr{F}_{m-1})$$

as the number *conjugate* to $x = A+B\sqrt{C}$ in $\mathscr{F}_m$. The numbers from $\mathscr{F}_{m-1}$, $\mathscr{F}_{m-2}$, ..., $\mathscr{F}$ are then conjugate to themselves in $\mathscr{F}_m$. One sees immediately the validity of the following rules

$$\begin{aligned} (x+y)^* &= x^*+y^*, \\ (xy)^* &= x^* \cdot y^*. \end{aligned} \tag{10}$$

If $f(x)$ is a polynomial with coefficients from $\mathscr{F}_{m-1}$, it follows from rules (10) that

$$(f(x))^* = f(x^*). \tag{11}$$

Since $0 = 0^*$, this means in particular that, besides $x_1 = a+b\sqrt{c}$

$$x_2 = x_1^* = a-b\sqrt{c}$$

is also a root of (9). The two roots are different from one another since otherwise $x_1 = x_2 = a$ would be true. However, by our hypothesis, a belongs to $\mathscr{F}_{m-1}$ and x_1 must not belong to $\mathscr{F}_{m-1}$.

Now the coefficients of a polynomial are, as is well-known,† elementary symmetric functions of the roots. In particular, for the coefficient a_1 of (9)

$$-a_1 = x_1+x_2+x_3.$$

Therefore, we have for the third root x_3 of (9)

$$x_3 = -(x_1+x_2)-a_1 = -a_1-2a.$$

Thus x_3 belongs to the field $\mathscr{F}_{m-1}$. Therefore, this number can be written as follows

$$x_3 = u+v\sqrt{w}, \quad u, v, w \in \mathscr{F}_{m-2}.$$

Hence it follows immediately that $x_3^* = u-v\sqrt{w}$ must also be a root of (9) which belongs to the field $\mathscr{F}_{m-1}$. It cannot be identically equal to x_1 or x_2 since these numbers do not belong to $\mathscr{F}_{m-1}$.

† See e.g. [IX 1].

Since (9), being an equation of third degree, has only three roots, $x_3^* = x_3$ must be true. However, this means that $w = 0$ and $x_3 \in \mathfrak{F}_{m-2}$. This argument can be repeated until one obtains the result that x_3 belongs to the field $\mathfrak{F}$ of coefficients itself. Hence Theorem 31 is proved.

In particular, it follows that

An equation of third degree with rational coefficients which can be solved by a constructible number † has also always a rational solution.

Hence we can immediately conclude that the *Delian problem* (*the doubling of the cube*) *is not solvable.* To prove this we have only to show that equation (4) cannot have a rational solution. Then by Theorem 31 there cannot be a solution from a field $\mathfrak{F}_m$.

Let us suppose that there is a solution $x = \frac{a}{b}$ of (4) with a and b relatively prime integers. Then we would have

$$a^3 = 2b^3.$$

However, this is impossible since a would be even and therefore b odd. Then a^3 would be divisible by 8 but $2b^3$ would not be.

Equation (3) for the trisection of the angle contains the parameter $\cos \alpha$. For $\alpha = \pi/2$ there exists the rational solution $y = 0$. The angle $\alpha = \pi/2$ can, of course, be divided into three equal parts using ruler and compasses since the angle $\pi/6$ can be constructed. On the other hand there does not exist a rational solution for $\alpha = \pi/3$. In this case we have

$$y^3 - 3y - 1 = 0. \tag{3'}$$

Let us suppose that (3') has a solution $y = \frac{a}{b}$ where a and b are relatively prime integers. Then we would have

$$a^3 - 3ab^2 - b^3 = 0.$$

But this means that each prime factor of a is also a prime factor of b and conversely. For a and b to be relatively prime, we would have $a = b = 1$. But $y = 1$ is not a solution of (3'). Thus we have

† A constructible number is a number from a field $\mathfrak{F}_m$ for which the field $\mathfrak{F}$ (see Theorem 30) is the field $\mathfrak{F}_0$ of rational numbers.

shown that (3′) does not have a rational solution. It follows from Theorem 31 that there does not exist a solution for equation (3′) from a field $\mathcal{F}_m$. However, this means that one cannot divide the angle $\frac{\pi}{3} = 60°$ into three equal parts using ruler and compasses. This establishes the fact that there cannot be a *general* method for trisecting *every* angle. Our results are summarised in

Theorem 32: *The Delian problem (doubling the cube using ruler and compasses) is not solvable. There exist angles which cannot be trisected using ruler and compasses.*

We now investigate the solvability of problem (**8d**). It led to an equation of fourth degree. As is well-known ([IX 1], vol. 2, p. 74) one can solve these equations by radicals. One first reduces the equation, using the Tschirnhaus transformation, to the form

$$y^4 + py^2 + qy + r = 0 \tag{12}$$

and then obtains the 4 solutions

$$y = \quad u+v+w \quad y = \quad u-v-w$$

$$y = -u+v-w \quad y = -u-v+w,$$

where u^2, v^2 and w^2 are the roots of the cubic equation

$$z^3 + \tfrac{1}{2}pz^2 + (\tfrac{1}{16}p^2 - \tfrac{1}{4}r)z - \tfrac{1}{64}q^2 = 0. \tag{13}$$

Equation (8) of problem (**8d**) is already in the form (12). The corresponding equation (13) in this case reads

$$z^3 - \tfrac{1}{4}\mu^2 z - \tfrac{1}{16}\mu^2 = 0. \tag{13′}$$

Then † (8) has a solution which can be constructed using ruler and compasses if and only if (13′) has a root which belongs to the field $\mathcal{F}(\mu)$ of coefficients of (13′). For $\mu = 1$, for example, if one assumes that there is a rational solution $z = \frac{a}{b}$ (where a and b are relatively prime integers) one obtains the equation

$$16a^3 - 4ab^2 - b^3 = 0.$$

† The reader should consider that this condition is not only sufficient but also necessary.

From this one argues:—Each prime factor A of a is also a prime factor of b. Since a and b are relatively prime, we have $a = 1$. For each prime factor B of b, B^2 must be a divisor of 16. Hence only the solutions $z = 1$, $z = \frac{1}{2}$ and $z = \frac{1}{4}$ are possible. Since none of these numbers is a solution of equation (13′) (for $\mu = 1$), there does not exist a rational solution of this equation. Hence we have established that problem (**8d**) has no *general* solution, since the construction is not possible for $\mu = 1$.†

Besides problem (**8d**) there exists a whole series of construction problems for the triangle which cannot (in general) be solved using ruler and compasses. We mention as examples the following (for (a) and (b) see [II 1] and for (c) see [IX 2, 3]):—

(a) To construct an isosceles triangle, given
  (α) s the half-perimeter ($2s = a+b+c$) and the incircle radius ρ,
  (β) s and the circumcircle radius r.

(b) To construct a right-angled triangle, given
  (α) the height h from the vertex opposite to the hypotenuse c and $c+a$,
  (β) h and $a+q$. (q is the projection of b on the hypotenuse c).

(c) To draw a triangle, given
  (α) a, b and ρ (incircle radius),
  (β) a, c and w_γ (angle bisector).

4. The Quinquesection of the Angle

Theorem 31 referring to equations of third degree cannot be used to say anything about the solution of equation (2). However, one can derive a similar statement for this equation of the quinquesection problem.

Equation (2) was constructed so that it is satisfied by $y_1 = 2\cos\frac{\alpha}{5}$. One can see immediately that the numbers

$$y_2 = 2\cos\left(\frac{\alpha}{5} + \frac{2\pi}{5}\right), \quad y_3 = 2\cos\left(\frac{\alpha}{5} - \frac{2\pi}{5}\right), \qquad (14)$$
$$y_4 = 2\cos\left(\frac{\alpha}{5} + \frac{4\pi}{5}\right), \quad y_4 = 2\cos\left(\frac{\alpha}{5} - \frac{4\pi}{5}\right)$$

† One easily sees that the case $\mu = 1$, therefore $J = r^2$, is geometrically realisable.

are also solutions of (2). On substituting

$$\frac{\alpha}{5}+\frac{2\pi k}{5}$$

into (1), one obtains

$$\cos 5\beta = 2\cos\left[5\left(\frac{\alpha}{5}+\frac{2\pi k}{5}\right)\right] = 2\cos\alpha =$$

$$\left[2\cos\left(\frac{\alpha}{5}+\frac{2\pi k}{5}\right)\right]^5 - 5\left[2\cos\left(\frac{\alpha}{5}+\frac{2\pi k}{5}\right)\right]^3 + 5\left[2\cos\left(\frac{\alpha}{5}+\frac{2\pi k}{5}\right)\right]$$

or

$$y_\nu^5 - 5y_\nu^3 + 5y_\nu - 2\cos\alpha = 0$$

for

$$\nu = 1, 2, 3, 4, 5.$$

The 5 roots $y_1, y_2, \ldots, y_5$ are all different for $\alpha \neq \pm\pi, \neq \pm 2\pi$. Now let us assume that the quinquesection of an angle $\alpha \neq \pi, \neq 2\pi$ using ruler and compasses is possible. Then (2) must have a root which belongs to a quadratic extension field $\mathcal{F}_m$ of the field of coefficients $\mathcal{F} = \mathcal{F}_0(\cos\alpha)$ of (2). Therefore let y_1 belong to a field $\mathcal{F}_m$ of this kind but not to the field $\mathcal{F}_{m-1}$.

Now we note that, from (14),

$$y_2 = 2\left(\frac{y_1}{2}\right)\cos 72° - 2\sqrt{1-\left(\frac{y_1}{2}\right)^2}\sin 72°$$

$$= \tfrac{1}{4}(\sqrt{5}-1)y_1 - 2\sqrt{1-\left(\frac{y_1}{2}\right)^2}\sqrt{\tfrac{5}{8}+\tfrac{1}{8}\sqrt{5}}. \quad (15)$$

We obtain similar representations for y_3, y_4 and y_5 from (14) by applying the addition theorem. Then the roots y_2, y_3, y_4 and y_5 belong in every case to the field $\mathcal{F}_{m+3}$ which one obtains from $\mathcal{F}_m$ by adjoining $\sqrt{5}$, $\sqrt{\frac{5}{8}+\frac{1}{8}\sqrt{5}}$ and $\sqrt{1-\left(\frac{y_1}{2}\right)^2}$. It can happen that one (or some) of these roots already belong to a field $\mathcal{F}_s$, where $s<m$. This would be possible for example if $\sqrt{1-\left(\frac{y_1}{2}\right)^2}$, $\sqrt{\frac{5}{8}+\frac{1}{8}\sqrt{5}}$ and $\sqrt{5}$ already belong to $\mathcal{F}_m$. In each case there exists for every number ν ($\nu = 1, 2, 3, 4, 5$) a number $N(\nu)$ such that

$$y_\nu \in \mathcal{F}_{N(\nu)}, \quad y_\nu \notin \mathcal{F}_{N(\nu)-1} \quad (16)$$

is true. An odd number of roots must belong to at least one number $N(\nu)$. Let s, say, be this number. Now, since (see p. 120) $y_\nu^* = a_\nu - b_\nu\sqrt{c_\nu}$ is also a root if $y_\nu = a_\nu + b_\nu\sqrt{c_\nu}$ is, this is only possible if $y_\nu = y_\nu^*$ for some ν. If $\mathfrak{F}_{N(\nu)}$ is not $\mathfrak{F}$ itself we would draw the conclusion that $y_{(\nu)}$ also belongs to $\mathfrak{F}_{N(\nu)-1}$. This contradicts (16) and therefore at least one root must belong to $\mathfrak{F}$. Hence we have

Theorem 31a: *An angle* α $(\alpha \neq \pi, \neq 2\pi)$ *can be divided into* 5 *equal parts using ruler and compasses if and only if equation* (2) *has a solution which belongs to the field* $\mathfrak{F} = \mathfrak{F}_0(\cos\alpha)$.

We have proved this theorem by a method which was suited particularly to equation (2). Instead one can derive Theorem 31a from the following general statement:—

If an equation $f(x) = 0$ *which is irreducible* † *in a field* $\mathfrak{F}$ *has at least one root from a field* $\mathfrak{F}_m$ ‡, *its degree must be a power of* 2 [IX 7, p. 71].

Since (2) is of degree 5, this equation can have a root from $\mathfrak{F}_m$ only if it is reducible in $\mathfrak{F}$. Hence one deduces the existence of a root from $\mathfrak{F}$.

Let us examine some of the special cases of quinquesection. For $\alpha = \pi/2$, from (2) we have

$$y^5 - 5y^3 + 5y = 0. \tag{2'}$$

On the other hand for $\alpha = \pi/3$ we have

$$y^5 - 5y^3 + 5y - 1 = 0. \tag{2''}$$

Both equations have a rational solution. For (2′) it is $y = 0$, for (2″) $y = 1$. In both cases the solution denoted by y_1 (cos 18° or cos 12° respectively) is irrational. The rational solution is $y_2 = \cos 90°$ in the first case and $y_3 = \cos(-60°)$ in the second case.

Therefore for the angles $\pi/2$ and $\pi/3$ quinquesection is possible. This is trivial if one considers that an angle of $18° = \frac{1}{2} \cdot 36°$ and $12° = 72° - 60°$ can be constructed using ruler and compasses. On the other hand there exists no rational solution of (2) for $\alpha = \cos^{-1}\frac{1}{3}$. The proof is similar to that for the examples of equations (3) and (8) given in the preceding section.

Now it is natural to consider the question whether it is the rule

† For the definition see e.g. [IX 1], [IX 4].
‡ See footnote ‡, p. 117.

or the exception to be able to solve the dissection problem. This question is answered in the following

Theorem 33: *The set $\mathcal{S}_5$ ($\mathcal{S}_3$) of angles $\alpha(0<\alpha<2\pi)$ for which the quinquesection (trisection) is possible using ruler and compasses is denumerable. The set $\mathcal{T}_5$ ($\mathcal{T}_3$) of angles for which quinquesection (trisection) is not possible has the same cardinality as the continuum.*

This is easily proved. If quinquesection of an angle α using ruler and compasses is possible, there exists a root x_ν of equation (2) which belongs to the field $\mathcal{F}_0(\cos\alpha)$. But this means that x_ν is a rational function of $\cos\alpha$, say of the form

$$x_\nu = \frac{A_0 + A_1\cos\alpha + \ldots + A_n\cos^n\alpha}{B_0 + B_1\cos\alpha + \ldots + B_m\cos^m\alpha}, \quad A_\nu,\ B_\nu \text{ integral.} \qquad (17)$$

If one substitutes (17) in (2) one obtains for $\cos\alpha$ an algebraic equation with integral coefficients. But this means that $\cos\alpha$ is an algebraic number. Thus for all α for which $\cos\alpha$ is transcendental quinquesection cannot be carried out. The same is true for trisection. Since the set of all algebraic numbers is denumerable, our theorem is proved.

In addition we note that the denumerable set $\mathcal{S}_5$ ($\mathcal{S}_3$) of numbers for which quinquesection (trisection) is possible in the interval $0<\alpha<2\pi$ is everywhere dense. For it is obvious that trisection is possible for all α of the form $\alpha = \frac{m}{2^n}\pi$, quinqesection for $\alpha = \frac{m}{3\,.\,2^n}\pi$ *in addition.*

5. The Uniqueness of the Construction

We know that there is no unique method for dividing a given angle into three or five equal parts. However, having obtained Theorem 33, it is natural to ask whether there exists a unique method of performing quinquesection or trisection at least *for a denumerable set* $\{\alpha_n\}$ of angles. In the case of trisection this question was first dealt with by Bieberbach [IX 7, p. 55] using a beautiful method from the theory of functions. We now take up the same question for quinquesection for which the corresponding statements are possible.

Let us suppose that there exists a unique method of performing quinquesection using ruler and compasses for a sequence $\{\alpha_n\}$ of angles. Then the uniqueness of the geometrical method would

correspond to a unique representation of

$$y_1^{(n)} = 2\cos\frac{\alpha_n}{5}$$

by constructible numbers. Then for the one root of (2), belonging to the field $\mathfrak{F}_0(\cos\alpha_n)$, by Theorem 31a (with $\cos\alpha_n$ instead of $\cos\alpha$) there must be a representation of the form

$$y_\nu^{(n)} = \frac{A_0 + A_1\cos\alpha_n + \ldots + A_k\cos^k\alpha_n}{B_0 + B_1\cos\alpha_n + \ldots + B_l\cos^l\alpha_n} = f(\cos\alpha_n),$$

where the A_κ and B_λ are rational numbers which are *independent of n*. Hence, for some k ($k = 0, \pm 1, \pm 2$), the equation

$$F(\alpha_n) = 2\cos\left(\frac{\alpha_n}{5} + \frac{2\pi k}{5}\right) - f(\cos\alpha_n) = 0$$

would be satisfied for an infinite set of numbers n. Now, since the magnitudes of the angles α_n have at least one limit point between 0 and 2π, it follows by a well-known theorem from the theory of functions that the function

$$F(z) = 2\cos\left(\frac{z}{5} + \frac{2k\pi}{5}\right) - f(z) \tag{18}$$

is identically zero. But $f(z)$ has period 2π. Hence

$$0 = 2\cos\left(\frac{z}{5} + \frac{2k\pi}{5}\right) - f(z) = 2\cos\left(\frac{z}{5} + \frac{2\pi(k+1)}{5}\right) - f(z),$$

therefore

$$\cos\left(\frac{z}{5} + \frac{2k\pi}{5}\right) = \cos\left(\frac{z}{5} + \frac{2\pi(k+1)}{5}\right)$$

identically in z. But this is not true; therefore there cannot be a unique method of construction for a sequence $\{\alpha_n\}$ of angles.

However, as BIEBERBACH has shown, it is always possible to find a unique method for a *finite* number of angles.

Chapter X

CONSTRUCTIONS ON THE SPHERE

1. Stereographic Projection

Construction problems on the sphere (e.g. problem (9), p. 3) can be reduced to problems of plane geometry using stereographic projection. The properties of this projection are usually described in detail in text-books on elementary and analytical geometry and also in the introductory chapters of books on function theory.† Therefore we can limit ourselves here to collecting the rules of stereographic projection which are important for our purposes.

We call an arbitrary great circle on the unit sphere the " equator " $\mathscr{E}$ and the points of the sphere on the diameter normal to the plane of $\mathscr{E}$ the "north and south poles" (N and S). The stereographic projection used is the central projection of the points of the sphere from the centre N onto the plane p of the equator $\mathscr{E}$. It has the following properties:†—

1. There is exactly one image point in the plane p corresponding to each point on the sphere other than N.
2. Circles on the sphere are mapped into circles or straight lines on the plane p.
3. Angles are invariant under the projection.
4. If P and Q are diametrically opposite points on the sphere and P' and Q' are their stereographic images in the plane,
$$S'P' \,.\, S'Q' = 1 \qquad (1)$$
is true where S' is the image of the south pole and therefore the centre of the sphere. (1) is easily proved by applying the altitude theorem to fig. 41.
5. If ξ, η and ζ are the coordinates of the point P on the sphere and x and y are those of the image point P' in the plane p,
$$x = \frac{\xi}{1-\zeta},\ y = \frac{\eta}{1-\zeta} \qquad (2)$$

† See e.g. [II 4] or [VI 2].

and

$$\xi = \frac{2x}{1+x^2+y^2}, \quad \eta = \frac{2y}{1+x^2+y^2}, \quad \zeta = \frac{x^2+y^2-1}{x^2+y^2+1}. \qquad (2')$$

Rule (1) can be used to perform constructions in the stereographic image in the plane. In order to draw, say, the image of the spherical quadrilateral mentioned in problem (9), we place the centre of this (regular) quadrilateral at the south pole of the sphere for simplicity. Then orthogonal great circles g_1 and g_2 through S are mapped into perpendicular straight lines g_1' and g_2' through S' and each " circle of latitude " l of the sphere becomes a circle with centre S' by the stereographic projection. Then the great circle arcs AB, BC, CD

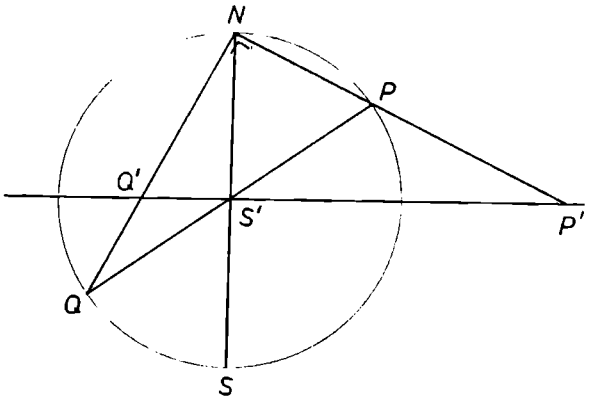


FIG. 41

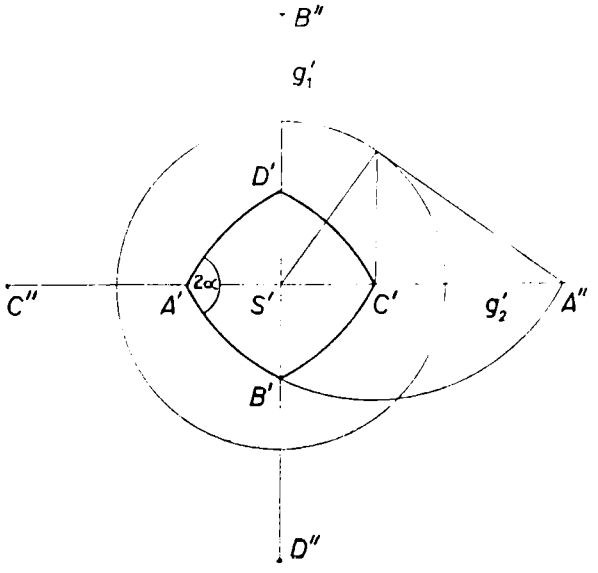


FIG. 42

and DA, determined by the points of intersection A, B, C and D of g_1 and g_2 with l, form a regular quadrilateral. In order to draw the images of the great circles through the image points A', B', C' and D', from (1), we only require *to reflect* these points *in S' and the unit circle.* Then the images A'', B'', C'' and D'' obtained in this way determine the images of the great circle arcs to be drawn in the plane p (fig. 42): the image of the great circle through A and B for example must pass through A'' (and B'') as well as A' and B'. Since angles are invariant under stereographic projection, the angle 2α in fig. 42 is equal to the (obtuse) angle of the quadrilateral on the sphere itself.†

2. Constructions with Spherical Compasses and Spherical Ruler

To examine the possibilities of constructions using spherical compasses and spherical ruler (problem (9), p. 3) in the stereographic image, it is useful to specify a rectangular coordinate system on the sphere. Then by (2') and (2) we can calculate the coordinates of the original points from those of the images, and conversely. However, we must first test whether it is always possible to construct three mutually orthogonal great circles from $n \geqq 2$ given points.

Let us suppose that two points P and Q are given on the sphere. If these two points lie on a diameter, one cannot construct any other point. We do not get much further if the distance between the two points is exactly $\frac{2}{3}\pi$. The circles with centre A and radius ‡ AB and with centre B and radius BA touch each other in a point C of the great circle AB and the distance of C from A and B is $\frac{2}{3}\pi$. In this case it is not possible to construct other points.§ We exclude these two cases in the following.

† Remember that figs. 10 and 15 are also stereographic images.

‡ The radius $AB = r$ is the arc AB measured on the sphere, the span of the compasses is the chord AB.

§ On page 4 it was mentioned that the spherical ruler is equivalent to the use of ordinary compasses provided one is given the span $\sqrt{2}$ for the great circle. First one determines the pole of the great circle and with this as centre draws a circle with span $\sqrt{2}$. One should note that drawing circles with the span $\sqrt{2}$ is *not in itself* a permissible construction step. This possibility was mentioned only to show how one can realise the " spherical ruler " in practice. Only the circles (with a span $< \sqrt{2}$) and the " lines " drawn by some means or other through two points on the sphere are construction elements. From this standpoint it is not possible to construct the pole of the great circle ABC where the (spherical) distances between the points A, B and C are $\frac{2}{3}\pi$.

If (at least) two points P and Q are given whose distance apart is not equal to π or $\frac{2}{3}\pi$, one can always construct three mutually orthogonal great circles which then determine a coordinate system on the sphere. One of these is the great circle PQ which we call the " equator ". If the arc PQ is less than $\frac{2}{3}\pi$, the circles with centre P and radius PQ and with centre Q and radius QP intersect each other and the great circle, passing through the points of intersection, goes through N and S, bisects the arc PQ, and is orthogonal to the equator

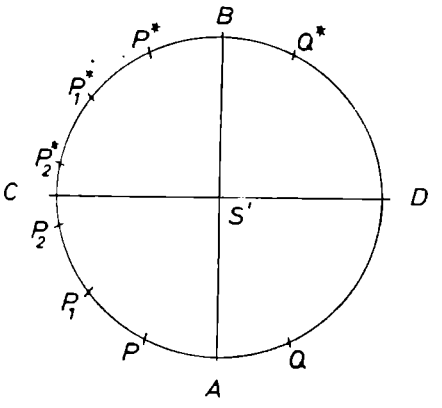


FIG. 43

$\mathscr{E}$. Let the points of intersection with $\mathscr{E}$ be A and B (fig. 43). Thus we have a pair of orthogonal great circles. To find the great circle orthogonal to these two, we draw the circle with centre B and radius AP which intersects $\mathscr{E}$ at P^* and Q^*. Then the perpendicular bisector of the arc of the equator PP^* (or QQ^*) is orthogonal to AB and PQ. However, since the arc PP^* is greater than $\frac{2}{3}\pi$ we certainly cannot get this third great circle by the method we used to obtain the second. The procedure here is to measure off the radius AP from P and P^* on the equator repeatedly and so obtain the pairs of points $P_1P_1^*$, $P_2P_2^*$, ... After a finite number of steps one obtains a pair $P_\nu P_\nu^*$ whose spherical distance apart is smaller than $\frac{2}{3}\pi$. Then the bisector of *this* arc of the equator can be constructed in the usual way and is the third great circle which intersects the equator orthogonally in C and D and goes through N and S.

Then the axes AB, CD and NS determine a rectangular coordinate system which we can use to represent the stereographic projection analytically by (2) or (2′).

If the spherical distance between the two given points P and Q is greater than $\frac{2}{3}\pi$, the method must be altered and we leave the reader to supply the details of this. The coordinates ξ, η and ζ

appearing in formulae (2) and (2′) can be interpreted as half the span of the compasses between two given or constructed points. In fact, if we let P be a point on the sphere with coordinates ξ, η, ζ, we can construct from P and the points determining the coordinate system † the reflected point $\bar{P}$ with coordinates $-\xi, \eta, \zeta$ using spherical compasses and spherical ruler. Then the span of the compasses between P and $\bar{P}$ is exactly 2ξ. Similarly, 2η and 2ζ can be interpreted as the spans of the compasses between constructed points.

Now, corresponding to each construction step on the sphere there is a well-defined operation using ruler and compasses in the stereographic image. Conversely, each step in terms of circles and straight lines in the image plane has an equivalent on the sphere since, of course, the circles and lines of the plane are mapped by the inverse of the stereographic projection into circles or great circles on the sphere. Since by (2) and (2′) the relations between (x, y) and (ξ, η, ζ) are rational, the possibilities for constructions using spherical ruler and spherical compasses can be described from Theorem 30 as follows:

Theorem 34: *Let $P_1, P_2, \ldots, P_n$ be n $(n \geqq 2)$ given points ‡ on the sphere with coordinates $(\xi_\nu, \eta_\nu, \zeta_\nu)$ $(\nu = 1, 2, \ldots, n)$. Let $\mathfrak{F}$ be the number field obtained from the field $\mathfrak{F}_0$ of rational numbers by adjoining the coordinates $\xi_\nu, \eta_\nu, \zeta_\nu$, $(\nu = 1, 2, \ldots, n)$. Then one can construct on the sphere, using spherical compasses and spherical ruler, just the points whose coordinates belong to a quadratic extension § field $\mathfrak{F}_m$ of $\mathfrak{F}$.*

Since there is the simple relation $s = \sin \frac{r}{2}$ between the span s of the compasses and the corresponding great circle arc r, Theorem 34 can also be formulated as follows [IX 7]:

Theorem 34a: *Given m great circle arcs, one can construct, using spherical compasses and spherical ruler, just the great circle arcs whose trigonometric functions are numbers constructible from the trigonometric functions of the given great circle arcs.*

† These are the points denoted above by A, B, C, D, S and N.

‡ We exclude the cases when only two points are given and the spherical distance between them is π or $\frac{2}{3}\pi$.

§ See footnote ‡, p. 117.

3. Squaring the Circle on a Sphere

Having dealt with these preliminaries we can go on to examine the squaring problem on the sphere, problem (9). We transfer the centre of the given spherical cap with radius ρ to the south pole of the sphere and consider as given the centre S $(0,0,-1)$ and a point P $(\sin\rho, 0, -\cos\rho)$ on the circumference.

The area of this spherical cap is

$$F_1 = 2\pi(1-\cos\rho). \tag{3}$$

On the other hand the area of a regular spherical quadrilateral with angle 2α (fig. 42) is

$$F_2 = 8\alpha - 2\pi. \tag{4}$$

Now we have the relation

$$\cot\alpha \,.\, \cot\frac{\pi}{4} = \cot\alpha = \cos r \tag{5}$$

between the circumcircle radius r of the regular quadrilateral and the angle α. Then, from $F_1 = F_2$, it follows that

$$\frac{\pi}{2} - \alpha = \frac{\pi}{4}\cos\rho \tag{6}$$

or, from (5),

$$\cos r = \tan\left(\frac{\pi}{4}\cos\rho\right). \tag{6'}$$

Now we can certainly draw the required quadrilateral if its circumcircle is known. Therefore we can also formulate our problem as follows:

We are given the south pole and a point P with coordinates $(\sin\rho, 0, -\cos\rho)$. *It is required to find a point Q with coordinates* $(\sin r, 0, -\cos r)$ *so that relation* (6′) *between r and* ρ *is satisfied.*

By (2) the stereographic images P' and Q' have the coordinates

$$x = \frac{\sin\rho}{1+\cos\rho}, \quad y = 0 \tag{7}$$

and

$$x_1 = \frac{\sin r}{1+\cos r}, \quad y_1 = 0. \tag{7'}$$

Then the relation

$$x_1^2 = \frac{1-\tan\left(\frac{\pi}{4}\frac{1-x^2}{1+x^2}\right)}{1+\tan\left(\frac{\pi}{4}\frac{1-x^2}{1+x^2}\right)} \tag{8}$$

between the coordinates of the stereographic image corresponds to formula (6'). Therefore the construction on the sphere is practicable if and only if x_1 in the plane can be constructed from given x using ruler and compasses. This can be done if and only if the angle

$$\beta = \frac{\pi}{4}\frac{1-x^2}{1+x^2} = \frac{\pi}{4}\cos\rho$$

can be constructed using ruler and compasses. One sees immediately that, for example, the conditions

$$\cos\rho = 2^{2-n}, \quad \cos\rho = 3^{-1}\,.\,2^{3-n} \quad (n = 2, 3, \ldots)$$

are sufficient for this.

But the construction is also possible for some cases when $\cos\beta$ is irrational. Thus for example $\beta = \tan^{-1}\frac{4}{3}$ can, of course, be constructed although $\tan^{-1}\frac{4}{3} = \cos^{-1}\frac{3}{5}$ *is not a rational multiple of* π. We can prove this fact using a lemma which is given in the last section of this chapter.

Therefore, although the squaring problem in the plane † is *unsolvable, the corresponding problem on the sphere can be solved in some cases.*

Theorem 35: *It is possible to construct a regular quadrilateral equal in surface area to a spherical cap of radius* ρ *using spherical compasses and spherical ruler if and only if* $\tan\left(\frac{\pi}{4}\cos\rho\right)$ *belongs to a quadratic extension* ‡ *field* $\mathfrak{F}_m$ *of the field* $\mathfrak{F} = \mathfrak{F}_0(\cos\rho)$ *where* $\mathfrak{F}$ *is the field obtained from the field* $\mathfrak{F}_0$ *of rational numbers by adjoining* $\cos\rho$.

There exists an analogue to this theorem in non-Euclidean (hyperbolic) geometry.§

One can also state the converse of problem (9):

(9a) *If a regular quadrilateral which is inscribed in a spherical cap of spherical radius r and which has angle* 2α *is given, a spherical cap of equal surface area is to be constructed using spherical compasses and spherical ruler.*

To investigate the solvability of this problem it is convenient to

† See e.g. [IX 7], [X 2], [X 3], [X 4].

‡ See footnote ‡, p. 117.

§ See [X 1]. However, the condition formulated there for the practicability of the construction is not necessary, but only sufficient.

start from the angle α of the given quadrilateral. If one puts $\alpha = \frac{\pi\omega}{4}$, from (6) we have

$$\cos\rho = 2-\omega, \quad \sin\rho = \sqrt{4\omega-\omega^2-3}.$$

Hence $\cos\rho$, $\sin\rho$ and thence the x-coordinate of P' by (7) can be constructed if and only if ω belongs to a quadratic extension field $\mathcal{F}_m^*$ of the field $\mathcal{F}^* = \mathcal{F}_0(\cos r)$.

4. A Lemma

We have still to prove that λ in

$$\beta = \lambda\pi = \tan^{-1}\tfrac{4}{3} = \cos^{-1}\tfrac{3}{5}$$

is irrational. This can be done using the following general

Lemma: *The vertex angles of a Pythagorean triangle are not rational multiples of* π.

To prove this theorem let us suppose that the Pythagorean numbers u, g and r are relatively prime: i.e.

$$u^2+g^2 = r^2, \quad (u, g, r) = 1.$$

As is well-known, one of the side lengths of the legs of a right-angled triangle must be even and one odd. Therefore let us assume that u is odd and g even. We want to show that λ in

$$\beta = \lambda\pi = \tan^{-1}\frac{g}{u} = \cos^{-1}\frac{u}{r} \tag{9}$$

is irrational. Let us assume that $\lambda = \frac{m}{n}$, $(m, n) = 1$.

1st case: n is odd.

From the formula (see VI (7))

$$\sin n\beta = \binom{n}{1}\cos^{n-1}\beta\sin\beta - \binom{n}{3}\cos^{n-3}\beta\sin^3\beta + \ldots + (-1)^{\frac{n-1}{2}}\sin^n\beta \tag{10}$$

for n odd we obtain in our case

$$0 = \binom{n}{1}\left(\frac{u}{r}\right)^{n-1}\left(\frac{g}{r}\right) - \binom{n}{3}\left(\frac{u}{r}\right)^{n-3}\left(\frac{g}{r}\right)^3 + - \ldots + (-1)^{\frac{n-1}{2}}\left(\frac{g}{r}\right)^n.$$

Hence it follows that

$$nu^{n-1} = \binom{n}{3} u^{n-3} g^2 - + \ldots + (-1)^{\frac{n+1}{2}} g^{n-1} = G$$

with G an even number. However, this is impossible since $n \,.\, u^{n-1}$ is odd.

2nd case: n is even: $n = 2^\rho U$, where U is odd.

Hence we put †

$$\gamma = 2^\rho \,.\, \beta = \frac{m\pi}{U} \tag{11}$$

and then note that $\sin\gamma$ and $\cos\gamma$ may be written in the form

$$\sin\gamma = \frac{g_\rho}{r^{2\rho}}, \quad \cos\gamma = \frac{u_\rho}{r^{2\rho}} \tag{12}$$

where g_ρ is even and u_ρ is odd. The validity of (12) is proved from (9) using the addition theorem and induction. We now apply (10) to $\sin U\gamma$ and obtain

$$0 = \binom{U}{1}\left(\frac{u_\rho}{r^{2\rho}}\right)^{U-1}\left(\frac{g_\rho}{r^{2\rho}}\right) - \binom{U}{3}\left(\frac{u_\rho}{r^{2\rho}}\right)^{U-3}\left(\frac{g_\rho}{r^{2\rho}}\right)^3 + - \ldots + (-1)^{\frac{U-1}{2}}\left(\frac{g_\rho}{r^{2\rho}}\right)^U$$

or

$$U \,.\, u_\rho^{U-1} = \binom{U}{3} u_\rho^{U-3}\, g_\rho^2 - + \ldots + (-1)^{\frac{U+1}{2}} g_\rho^{U-1} = G^*$$

where G^* is an even number and U and u_ρ are odd. From this contradiction it follows that λ cannot be rational in (9). Then μ in $\mu\pi = \dfrac{\pi}{2} - \beta$ is not rational and our lemma has been proved.

† $U \neq 1$ since $U = 1$ leads to $\sin\gamma = \sin m\pi = 0$. Obviously $\sin 2^\rho\beta$ is different from 0 for all ρ.

Chapter XI

PROBLEMS IN THE GEOMETRY OF SETS

1. A New Definition of Equivalence by Dissection

In elementary geometry the isosceles triangle ABC in fig. 44 is equivalent by dissection (p. 2) to the rectangle $ADCE$: the triangle

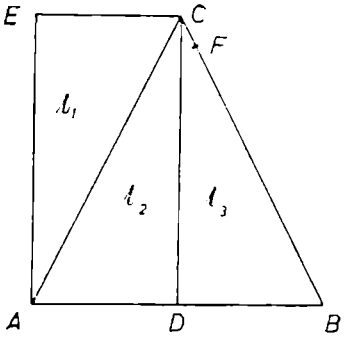


Fig. 44

ACD is congruent to itself and the triangle ACE to triangle BCD. Here we want to explain that the equivalence by dissection

$$ABC \stackrel{d}{=} ADCE$$

does not immediately establish a one-to-one correspondence between the sets of points determined by the triangle and the rectangle. Let r, say, be the set of points which lie in the interior of the rectangle $ADCE$, t the corresponding set for the triangle ABC and let t_1, t_2 and t_3 be the sets of interior points of ΔACE, ΔACD and ΔBCD. Then to the dissection

$$ACD + ACE = ACD + BCD$$

there corresponds a one-to-one correspondence of t_1 onto t_3 and of t_2 onto itself. The first of these two correspondences is realised by the congruence mapping which brings the two figures into superposition.

However, no equivalent correspondence is established between i and $\mathfrak{l}$. The union of the sets $\mathfrak{l}_1$ and $\mathfrak{l}_2$ is a proper subset of i since the interior points of the line AC belong to i but not to $\mathfrak{l}_1+\mathfrak{l}_2$. Likewise the interior points of CD belong to $\mathfrak{l}$ but not to $\mathfrak{l}_2+\mathfrak{l}_3$. It is not any simpler to consider, instead of i, $\mathfrak{l}$ and $\mathfrak{l}_\nu$ ($\nu = 1, 2, 3$), the corresponding closed sets $\mathfrak{R}$, $\mathfrak{T}$, $\mathfrak{T}_\nu$ ($\nu = 1, 2, 3$).

Nevertheless it is possible to translate the equivalence by dissection in elementary geometry into set theory. For this purpose we give the definition

Two point sets $\mathfrak{P}$ and $\mathfrak{Q}$ (of an n-dimensional Euclidean space) are said to be "equivalent by dissection into n parts" if it is possible to decompose $\mathfrak{P}$ and $\mathfrak{Q}$ into n mutually disjoint sets $\mathfrak{P}_\nu$ and $\mathfrak{Q}_\nu$ respectively ($\nu = 1, 2, \ldots, n$) so that the subsets $\mathfrak{P}_\nu$ and $\mathfrak{Q}_\nu$ with the same suffix are congruent†:

$$\mathfrak{P} = \mathfrak{P}_1 + \mathfrak{P}_2 + \ldots + \mathfrak{P}_n, \quad \mathfrak{P}_\nu \mathfrak{P}_\mu = 0 \quad (\nu \neq \mu),$$

$$\mathfrak{Q} = \mathfrak{Q}_1 + \mathfrak{Q}_2 + \ldots + \mathfrak{Q}_n, \quad \mathfrak{Q}_\nu \mathfrak{Q}_\mu = 0 \quad (\nu \neq \mu),$$

$$\mathfrak{P}_\nu \equiv \mathfrak{Q}_\nu, \quad \nu = 1, 2, \ldots, n.$$

Then we write $\mathfrak{P} \overset{n}{\equiv} \mathfrak{Q}$. If the number n of the parts is not known (or there is no particular point in knowing this number), one can say simply that the sets are finitely equivalent: $\mathfrak{P} \overset{(f)}{\equiv} \mathfrak{Q}$, where two sets are said to be *congruent* if it is possible to transform the one into the other by congruence mappings (displacements, rotations and reflections).

This equivalence by dissection into n parts must certainly be distinguished from the elementary equivalence by dissection of polygons or polyhedra introduced earlier. In the new definition any point sets of the considered space, not only polygons and polyhedra, are permitted for the sets *and the subsets*. On the other hand we require here for equivalence a *unique correspondence of the points and a dissection into mutually disjoint sets*. In the dissection of polygons into sub-polygons it was required only that the sub-polygons should have no *interior* points in common.

We now show that the rectangle $ADCE$ and the triangle ABC of fig. 44 are also equivalent by finite dissection as just defined, where we now mean the rectangle or the triangle to be taken as the

† On the addition and subtraction of sets see e.g. [VI 5] or another text-book on set theory.

set of points which belong to the interior or to the boundary of the figure. In the following the lines AC, BC and so on are enclosed in brackets so that we can indicate whether we are referring to the open, half-open or closed set. Thus $[AC]$ (in square brackets) means the line AC including the end points, (AC) is the open line and $[AC)$ the " half-open " line (including A, excluding C).

We now try to obtain the correspondence of the sets of points under consideration as follows: The (closed) triangle ACD ($\mathscr{T}_2$) corresponds to itself, the open triangle ACE (ℓ_1 in fig. 44) to the open triangle BCD (ℓ_3):

$$\mathscr{T}_2 \equiv \mathscr{T}_2, \ell_1 \equiv \ell_3.$$

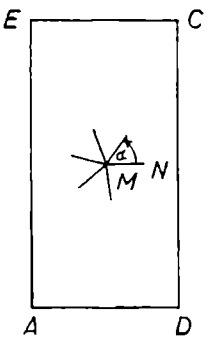


FIG. 45

Now we have still to dispose of some lines. Let F be the point on $[BC]$ for which the congruence $[AE] \equiv [BF]$ is satisfied. Then we can decompose the sets $\mathscr{R}$ (the closed rectangle $ADCE$) and $\mathscr{T}$ (the closed triangle ABC) as follows:—

$$\mathscr{R} = \mathscr{T}_2 + \ell_1 + (EC) + (AE],$$

$$\mathscr{T} = \mathscr{T}_2 + \ell_3 + (DB) + (FB] + (CF],$$

where the congruence relations hold good:

$$\mathscr{T}_2 \equiv \mathscr{T}_2, \quad \ell_1 \equiv \ell_3, \quad (EC) \equiv (DB), (AE] \equiv (FB].$$

We can summarise this result as follows:—

$$\mathscr{T} \overset{5}{\equiv} (CF] + \mathscr{R}. \tag{1}$$

This means that the line $(CF]$ is simply left over! However, we will manage to accommodate it in the rectangle.

For this purpose we consider the following. Let M be the centre of the rectangle $ADCE$ (fig. 45) and $(MN]$ a (half-open) line parallel to EC and with length less than $\frac{1}{2}EC$. We denote this line by σ and by $\sigma(\nu\alpha)$ (ν integral and positive) its image obtained by rotation about M through the angle $\nu\alpha$. Now we choose an arbitrary angle α which is not a rational multiple of π and define the set $\mathscr{S}$ as the union of sets of infinitely many $\sigma(\nu\alpha)$ ($\nu = 1, 2, 3, \ldots$):

$$\mathscr{S} = \sigma + \sigma(\alpha) + \sigma(2\alpha) + \ldots.$$

By rotating through the angle $-\alpha$ this set is transformed into

$$\mathscr{S}(-\alpha) = \sigma(-\alpha) + \sigma + \sigma(\alpha) + \ldots = \mathscr{S} + \sigma(-\alpha).$$

Therefore we have

$$\mathcal{S} = \mathcal{S} + \sigma(-\alpha). \tag{2}$$

Now, if $AD \leqq CD$ in the rectangle $\mathcal{R}$, CF is less than $\frac{1}{2}EC$ as one can easily check. We now choose the line σ to be congruent to $(CF]$ so, by (2), we have

$$\mathcal{S} \overset{2}{\equiv} \mathcal{S} + (CF]. \tag{3}$$

Now $\mathcal{R}$ is decomposed into $\mathcal{S}$ and the complementary set $\mathcal{R}_1$. Then from (3) we have

$$\mathcal{R} = \mathcal{R}_1 + \mathcal{S} \overset{3}{\equiv} \mathcal{R}_1 + \mathcal{S} + (CF] = \mathcal{R} + (CF]. \tag{4}$$

A transitivity relation similar to that for equivalence by dissection in elementary geometry is valid for finite equivalence:

Lemma A: *If two sets are finitely equivalent to a third, they are finitely equivalent to each other.*

i.e. If

$$\mathcal{A} \overset{m}{\equiv} \mathcal{B}, \quad \mathcal{B} \overset{n}{\equiv} \mathcal{C},$$

then

$$\mathcal{A} \overset{m \cdot n}{\equiv} \mathcal{C}.$$

One can carry out the proof of this lemma similarly (by decomposition into subsets) to that of the corresponding statement for polygons. We omit it here.†

There follows now from this Lemma A by (1) and (4):—

Theorem 36: *Each rectangle $\mathcal{R}$ ($ADCE$ in fig. 44) is finitely equivalent to the isosceles triangle $\mathcal{T}$ (ABC) formed by its two right-angled sub-triangles.*

Thus we have obtained a first theorem in the geometry of sets analogous to a well-known statement of elementary geometry. However, one achieves much more in the geometry of sets than giving a new interpretation to theorems like this. In this theory there are statements which must appear paradoxical from the standpoint of elementary geometry. One of these is, for example, the

Theorem 37: *Each sphere in three dimensional space $\mathcal{R}_3$ is finitely equivalent to two disjoint spheres of equal radius.*

Perhaps it is necessary at this stage to emphasise that " finitely

† It is in SIERPINSKI [XI 3], p. 69.

equivalent " says essentially more than " equivalent " in the general set-theoretical sense. Here it is asserted that each sphere can be dissected into finitely many parts so that *two* spheres of equal radius can be formed *by rotations and parallel displacements* from these parts. Using SIERPINSKI's terminology we call *the decomposition of a set $\mathcal{M}$ into two disjoint sets $\mathcal{M}_1$ and $\mathcal{M}_2$ paradoxical if each of the two subsets $\mathcal{M}_\nu$ ($\nu = 1, 2$) is finitely equivalent to $\mathcal{M}$.* One can derive peculiar statements from the paradox of Theorem 37. For example the paradox that each sphere of $\mathcal{R}_3$ is finitely equivalent to each cube of $\mathcal{R}_3$ (Theorem 39, p. 154). In addition, we will show (Theorem 40) that there are no paradoxical dissections in $\mathcal{R}_1$. This suggests considering what the situation is in $\mathcal{R}_2$. We will not anticipate any other results of the following sections but just stress that problem (**10**), p. 4), the squaring problem in the geometry of sets, is still unsolved. In order to solve this range of problems one must undertake a somewhat complicated way of reasoning in the next section.

2. HAUSDORFF'S THEOREM

One is able to tackle paradoxical dissections in three-dimensional space by means of Hausdorff's theorem:

Theorem 38: *The surface area of the unit sphere $\mathcal{U}$ of $\mathcal{R}_3$ may be represented as the union of four mutually disjoint sets*:

$$\mathcal{U} = \mathcal{Q} + \mathcal{R} + \mathcal{S} + \mathcal{T}. \tag{5}$$

Of these sets $\mathcal{Q}$ is denumerable and the others have the same cardinal number as the continuum. These sets are related by the congruences

$$\mathcal{R} \equiv \mathcal{S}, \mathcal{S} \equiv \mathcal{T}, \mathcal{T} \equiv \mathcal{R}, \mathcal{R} \equiv \mathcal{S} + \mathcal{T}. \tag{5'}$$

In this case the congruence motions are rotations of the sphere.

We divide the proof of this theorem into several parts.

A. *Definitions of transformations*

We place a rectangular coordinate system at the centre of the sphere and call the positive u-axis a ray which starts from the origin of the xz-plane and which makes an angle $\frac{1}{2}\Theta$ with the positive z-axis and an angle $\frac{\pi}{2} - \frac{1}{2}\Theta$ with the positive x-axis. We will still be able to choose the value of Θ.

Now let

$$P' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \phi(P) = \phi \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{and} \quad P' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \psi(P) = \psi \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

be the rotations through 180° about the u-axis and through 120° about the z-axis respectively. These rotations ϕ and ψ can be represented as follows:

$$\begin{aligned} x' &= -x \cos \Theta \qquad + z \sin \Theta, \\ y' &= \qquad\qquad - y, \\ z' &= \quad x \sin \Theta \qquad + z \cos \Theta, \end{aligned} \tag{6}$$

and

$$\begin{aligned} x' &= -\tfrac{1}{2}x - \tfrac{1}{2}\sqrt{3}y, \\ y' &= \tfrac{1}{2}\sqrt{3}x - \tfrac{1}{2}y, \\ z' &= \qquad\qquad z, \end{aligned} \tag{6'}$$

respectively. Provided no misunderstandings are caused, we will indicate simply by juxtaposition of the factors ϕ and ψ the composition of motions $\phi(P)$ and $\psi(P)$ such as

$$\phi(\psi(\psi(\phi(\psi(P))))). \tag{7}$$

Therefore instead of (7) we would write $\phi\psi^2\phi\psi$. We agree that in products like this the operation at the right must always be performed first.

Now it is true that

$$\phi^2 = 1, \quad \psi^3 = 1, \tag{8}$$

and hence one can reduce operations like

$$\phi^5\psi^4\phi^2\psi; \ \phi^{13}\psi^4\phi$$

to

$$\phi\psi^2; \ \phi\psi\phi.$$

Now we assert that *one can show, by suitable choice of the angle* Θ, *that differently written reduced transformations are also geometrically different.*

According to our convention reduced transformations are

represented in one of the following forms:—

$$\alpha = \psi^{m_1}\phi\psi^{m_2}\phi\ldots\psi^{m_n}\phi, \quad \beta = \phi\psi^{m_1}\phi\psi^{m_2}\ldots\phi\psi^{m_n},$$
$$\gamma = \phi\psi^{m_1}\phi\psi^{m_2}\ldots\psi^{m_n}\phi, \quad \delta = \psi^{m_1}\phi\psi^{m_2}\phi\ldots\phi\psi^{m_n}, \tag{9}$$

where the exponents m_ν are the numbers 1 or 2. Now if, for example,

$$\phi\psi^2\phi\psi = \psi\phi,$$

it would follow that

$$\phi\psi^2\phi\psi\phi\psi^2 = \psi\phi^2\psi^2 = \psi^3 = 1.$$

Obviously it is sufficient for the proof of this assertion to verify that none of the products just denoted by α, β, γ or δ can be equal to 1. We show first of all that, for suitable choice of Θ, a rotation of type α can never be equal to 1. These rotations are composed of rotations of the forms $\psi\phi$ and $\psi^2\phi$. Let us calculate first $\psi\phi$ from (6) and (6′). By combining the two transformations one obtains (most simply by multiplying out the matrices) the representation

$$\begin{aligned} x'' &= \tfrac{1}{2}\cos\Theta \,.\, x + \tfrac{1}{2}\sqrt{3} \,.\, y - \tfrac{1}{2}\sin\Theta \,.\, z, \\ y'' &= -\tfrac{1}{2}\sqrt{3}\cos\Theta \,.\, x + \tfrac{1}{2}y + \tfrac{1}{2}\sqrt{3}\sin\Theta \,.\, z, \\ z'' &= \sin\Theta \,.\, x \qquad\qquad\qquad + \cos\Theta \,.\, z \end{aligned} \tag{10}$$

for the transformation

$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \psi\phi \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Since ψ^2 indicates only the rotation through 240° or $-120°$ about the z-axis, one obtains the representation of this rotation from (10) on replacing $\sqrt{3}$ always by $-\sqrt{3}$.

The image of the point $M_0 = (0, 0, 1)$ under the operation $\psi\phi$ is, from (10), $M_1 = (-\frac{1}{2}\sin\Theta, \frac{1}{2}\sqrt{3}\sin\Theta, \cos\Theta)$. If one applies to M_1 the motion $\psi\phi$, one obtains

$$M_2 = (\sin\Theta \,.\, f_1(\cos\Theta), \sin\Theta \,.\, f_2(\cos\Theta), f_3(\cos\Theta)),$$

where $f_\nu(v)$ is a polynomial in v. It is not necessary for our purposes to evaluate these functions exactly. Then the repeated application of these motions leads to

$$M_n = (\sin\Theta \,.\, f_1^{(n)}(\cos\Theta), \sin\Theta \,.\, f_2^{(n)}(\cos\Theta), f_3^{(n)}(\cos\Theta)).$$

In any case the third component z_n of the point M_n will be a polynomial of the type

$$z_n = a_0 \cos^n \Theta + a_1 \cos^{n-1} \Theta + \ldots + a_n.$$

The coefficients are of the form $a+b\sqrt{3}$, where a and b are rational numbers. Now if $\alpha = 1$, $M_n = M_0$ would have to be true and therefore z_n would also be equal to 1. Hence we have for $\xi = \cos \Theta$

$$A_0\xi^n + A_1\xi^{n-1} + \ldots + A_n = \sqrt{3}(B_0\xi^n + \ldots + B_n)$$

with rational numbers A_ν, B_ν. By squaring one obtains from this an equation which is satisfied only by algebraic numbers ξ. If one chooses for example $\xi = \cos \Theta = \pi/4$, this equation cannot be satisfied and therefore α is not equal to 1.

It only remains to show that no product in (9) of the form β, γ or δ can be equal to 1. Therefore let us suppose that $\beta = 1$. Then $\phi\beta\phi = \phi^2 = 1$ also and therefore

$$\phi[\phi\psi^{m_1}\ldots\psi^{m_n}]\phi = \psi^{m_1}\phi\ldots\psi^{m_n}\phi = 1.$$

However, this is a transformation of type α and we know from above that it cannot be equal to 1. Let us now consider the case $\gamma = 1$. According to this we would also have $\phi\gamma\phi = \phi^2 = 1$. For the special case $\gamma = \phi\psi^{m_1}\phi$ this means that $\psi^{m_1} = 1$ and this is certainly false for $m_1 = 1$ or $m_1 = 2$. In the general case $\phi\gamma\phi = 1$ leads to an equation of type $\delta = 1$. Therefore we will have completed our proof if we have shown that, for the choice made for Θ, $\delta = 1$ also leads to a contradiction.

To verify this we divide the elements of δ into two classes. Transformations of type δ, for which the sum of the exponents $m_1 + m_n \neq 3$, belong to the class C_1 and the others to the class C_2.

Now let $\delta = 1$ where δ belongs to C_1. Then

$$\delta = 1 = \psi^{-m_1}\delta\psi^{m_1} = \phi\psi^{m_2}\ldots\phi\psi^{m_1+m_n}.$$

However, for $m_1 + m_n \neq 3$ this is a transformation of the type β. It cannot be equal to 1. Finally there still remains the case

$$\delta = 1 \text{ and } \delta \in C_2.$$

Here

$$\phi\psi^{m_n}\delta\psi^{m_1}\phi = \phi\psi^3\phi = \phi^2 = 1$$

or

$$\phi\psi^{m_n}[\psi^{m_1}\phi\psi^{m_2}\ldots\phi\psi^{m_n}]\psi^{m_1}\phi = 1.$$

Because $m_1 + m_n = 3$ and $\phi^2 = \psi^3 = 1$ we get from this

$$\delta' = \psi^{m_2}\phi \ldots \psi_{m_n - 1} = 1.$$

This is again a transformation of the type δ but it has two terms fewer than δ. δ' cannot, as already proved, belong to the class C_1. Therefore the method just used can be continued to reduce factors until finally $\phi = 1$ or $\psi^{m_\nu} = 1$ ($m_\nu = 1$ or $m_\nu = 2$) remains left over and this is certainly false. Hence it is shown that all the different reduced transformations for the particular choice of Θ are really different geometrically.

The transformations constructed by combining the factors ϕ and ψ form a group $\mathcal{G}$. We continue the proof of Theorem 40 by proving a lemma about this group $\mathcal{G}$.

B. *Proof of a lemma*

Lemma B: *The elements of the group $\mathcal{G}$ can be divided into three disjoint classes A, B and C so that the following conditions are satisfied:*

1. *If $\sigma \in A$, $\phi\sigma \in B + C$ and conversely.*
2. *Each of the conditions $\sigma \in A$, $\psi\sigma \in B$, $\psi^2\sigma \in C$ implies the other two.*

In order to be able to carry out the above division of the elements of $\mathcal{G}$ into three classes we first of all perform another division of the group $\mathcal{G}$. Let G_n be the set of elements of $\mathcal{G}$ whose reduced representation contains no more than n factors ϕ, ψ, ψ^2 (ψ^2 is considered as *one* factor). Then obviously

$$G_1 = (1, \phi, \psi, \psi^2).$$

We now divide this subset into classes A, B and C:

$$1 \in A,\ \phi \in B,\ \psi \in B,\ \psi^2 \in C.$$

The statement of lemma B is true for this division. In fact, if $\sigma \in A$ (and $\sigma \in G_1$), $\sigma = 1$. Then, however, $\phi\sigma = \phi \in B \subset B + C$. Conversely, if $\phi\sigma \in B + C$, $\phi\sigma = \phi$ and therefore $\sigma = 1 \in A$.

If σ, $\psi\sigma$, $\psi^2\sigma$ are all from G_1 then only three values are possible for σ: $\sigma = 1$, $\sigma = \psi$, $\sigma = \psi^2$. Now let one of the relations

$$\sigma \in A,\ \psi\sigma \in B,\ \psi^2\sigma \in C \tag{11}$$

be correct. If $\sigma \in A$, $\sigma = 1$ and the second and third relations are satisfied. If $\psi\sigma \in B$, the first of the values possible for σ, namely $\sigma = 1$, is taken. Finally, if $\psi^2\sigma \in C$, it follows again that $\sigma = 1$.

The further division of $\mathcal{G}$ into the three classes A, B and C can be carried out now by *induction*. Let us assume that G_n is divided according to the lemma. Since $G_n \subset G_{n+1}$ it is sufficient to give the classification for $G_{n+1} - G_n$. These are just the elements of $\mathcal{G}$ which have exactly $n+1$ factors.

Now, for an element $\tau \in G_{n+1} - G_n$, there are three possibilities: τ must be one of the three types

$$\tau = \phi\sigma, \ \tau = \psi\sigma, \ \tau = \psi^2\sigma, \tag{12}$$

where σ is an element of G_n with exactly n factors. Now we give a table which, for each of the cases named in (12), allows one to read off the class to which τ belongs if that of σ is known:

σ	Class of τ		
	$\tau = \phi\sigma$	$\tau = \psi\sigma$	$\tau = \psi^2\sigma$
A	B	B	C
B	A	C	A
C	A	A	B

(13)

With this the division of $\mathcal{G}$ into three classes A, B and C by induction is completed. It will be shown that it has the property asserted in Lemma B. Before doing this let us note as an example how, starting from the division of G_1, the division of G_3 is obtained from table (13):

	G_1	$G_2 - G_1$	$G_3 - G_2$
A	1	$\phi\psi, \ \phi\psi^2, \ \psi^2\phi$	$\phi\psi\phi$
B	$\phi, \ \psi$		$\phi\psi^2\phi, \ \psi\phi\psi, \ \psi\phi\psi^2$
C	ψ^2	$\psi\phi$	$\psi^2\phi\psi, \ \psi^2\phi\psi^2$

Now we assume that all the statements of lemma B are correct for G_n and show that they are also correct for G_{n+1}.

First we deal with statement 1 of the lemma. Let σ and $\phi\sigma$ be elements from G_{n+1}. Then the following four cases are to be taken into consideration:

$$\text{(a) } \sigma \in G_n, \quad \phi\sigma \in G_n; \quad \text{(b) } \sigma \in G_n, \quad \phi\sigma \notin G_n;$$

$$\text{(c) } \sigma \notin G_n, \quad \phi\sigma \in G_n; \quad \text{(d) } \sigma \notin G_n, \quad \phi\sigma \notin G_n.$$

In case (a) statement 1 is correct by the induction hypothesis. In the case (b) σ consists of exactly n factors of which the first is ψ or ψ^2. If the first factor were ϕ then, since $\phi^2 = 1$, $\phi\sigma$ would have fewer factors than σ and so would have to belong to G_n. If $\sigma \in A$, (13) shows that $\phi\sigma \in B \subset B+C$. Conversely, if $\phi\sigma \in B+C$, $\phi\sigma \in B$ (from the column $\tau = \phi\sigma$ of (13)) and hence $\sigma \in A$.

In the case (c) the number of factors for $\phi\sigma = \rho$ is less than for σ. This is only possible if the following is true:

$$\sigma = \phi\rho, \quad \rho = \phi\sigma, \quad \rho = \psi^{n_1}\ldots.$$

Now table (13) must always be read so that the element with the *smaller* number of factors is inserted in the column under σ. Therefore in our case $\sigma = \phi\rho$ is to be used and in (13) τ is to be replaced by σ and σ by ρ. If one observes this one reads from the table that $\rho = \phi\sigma \in B+C$ follows from $\sigma = \phi\rho \in A$ and conversely.

Finally the case (d) is not possible. σ would have to consist of $n+1$ factors and ϕ cannot be the first. (Otherwise we would have, of course, $\phi\sigma \in G_n$). Therefore σ is of the form $\sigma = \psi^{n_1}\ldots$. But then $\phi\sigma$ would have $n+2$ factors and would not belong to G_{n+1}. Hence statement 1 of lemma B is proved.

For the proof of statement 2, suppose that $\sigma, \psi\sigma, \psi^2\sigma$ are elements of G_{n+1}. If all three already belong to G_n then our statement is correct by the induction hypothesis. It is impossible to have *none* of the elements belonging to G_n. For then σ would be of the form $\sigma = \phi\ldots$, and $\psi\sigma$ and $\psi^2\sigma$ would have $n+2$ factors. Therefore there remains only the case when some of the three elements $\sigma, \psi\sigma, \psi^2\sigma$ belong to G_n and others do not.

First of all we suppose $\sigma \in G_n$. Then σ must be of the form $\sigma = \phi\ldots$. Otherwise $\psi\sigma$ and $\psi^2\sigma$ would also belong to G_n. Then it follows from (13) and $\sigma \in A$ that

$$\psi\sigma \in B, \quad \psi^2\sigma \in C$$

and conversely.

Next suppose $\sigma \notin G_n$, $\psi\sigma \in G_n$. Here we again have the case that multiplying by ψ decreases the number of factors. Then σ is of the form $\psi^2\rho$, $\rho = \phi\ldots$, $\rho = \psi\sigma$. Using table (13) one can observe again that $\rho = \psi\sigma$ has fewer factors than σ. Then from $\sigma = \psi^2\rho \in A$ it follows from (13) that $\rho = \psi\sigma \in B$, $\psi^2\sigma = \psi\rho \in C$. Conversely, from $\psi\sigma \in B$ or $\psi^2\sigma \in C$ it follows that $\psi^2 . \psi\sigma = \sigma \in A$.

There only remains the case: $\sigma \notin G_n$, $\psi^2\sigma = \rho \in G_n$. Here σ is of the form $\sigma = \psi\rho$ and for $\sigma \in A$ it follows that $\psi^2\sigma = \rho \in C$ and $\psi\sigma = \psi^2\rho \in B$. From each of these last relations it follows again that $\sigma \in A$. Hence lemma B is completely proved.

C. *Application of the axiom of choice*

In the last part of our proof we shall use the axiom of choice of set theory (see for example [XI 6], p. 53). It can be formulated as follows:

> *Let $\mathfrak{M}$ be a set of non-empty, mutually disjoint sets $\mathfrak{M}_\alpha$. Then there is a set $\mathfrak{m}$ with the following property:—Each element m of $\mathfrak{m}$ is an element of a certain set $\mathfrak{M}_\alpha$ and $\mathfrak{m}$ has exactly one element in common with each of the sets $\mathfrak{M}_\alpha$.*

This axiom is a statement about existence only and in practice it is not always possible to determine the elements of this "set of choice".† Thus we shall have to satisfy ourselves in this case too that the existence of the set mentioned in Theorem 38 is proved by the axiom of choice; but no method of writing the division of the spherical surface into 4 subsets exists in practice.

First we denote by $\mathfrak{Q}$ the denumerable set mentioned in Theorem 38. This is the set of points of the spherical surface which remains fixed under one of the rotations of the group $\mathfrak{G}$. Of course, a pair of diametrically opposite points remains fixed under each spherical rotation, and so we have a denumerable set $\mathfrak{Q}$ of such fixed points.

If M is a point of $\mathfrak{U} - \mathfrak{Q}$ then $\sigma_1(M) \neq \sigma_2(M)$ for all rotations σ_1, σ_2 ($\sigma_1 \neq \sigma_2$) of our group $\mathfrak{G}$. For, if $\sigma_1(M) = \sigma_2(M)$, it would follow that $\sigma_1^{-1}\sigma_2(M) = 1$. But $\sigma_1^{-1}\sigma_2$ is also an element of $\mathfrak{G}$ and M would be a fixed point for $\sigma_1^{-1}\sigma_2$. Since M does not belong to $\mathfrak{Q}$, this is impossible.

Now we denote by $\mathfrak{H}(M)$, for $M \in \mathfrak{U} - \mathfrak{Q}$, the set of all points $\sigma(M)$ which arise when σ runs through the whole group $\mathfrak{G}$. From

† On the difficulty of the "actual" evaluation see for example [VI 10], pp. 44 *et seq.*

this definition it follows immediately that, if $M \in \mathfrak{U} - \mathfrak{L}$, $N \in \mathfrak{U} - \mathfrak{L}$, $M \neq N$, the sets $\mathfrak{H}(M)$ and $\mathfrak{H}(N)$ are either disjoint or identical. Of course, if $\sigma_1(M) = \sigma_2(N)$, it follows that $M = \sigma_1^{-1}\sigma_2(N) = \tau(N)$. This means that M belongs to the set $\mathfrak{H}(N)$, and, since $\sigma(M) = \sigma\tau(N)$, $\mathfrak{H}(M)$ and $\mathfrak{H}(N)$ are identical. Therefore, in this way, $\mathfrak{U} - \mathfrak{L}$ is divided into a collection (of cardinality equal to that of the continuum) of classes of denumerable sets $\mathfrak{H}(M)$. By the axiom of choice there exists a set $\mathfrak{X}$ of " representatives ", i.e. a set which contains *exactly one element* from each of the sets $\mathfrak{H}(M)$.

Now, using this set of choice $\mathfrak{X}$, we can define the sets $\mathfrak{R}$, $\mathfrak{S}$ and $\mathfrak{T}$ of Theorem 38. First we agree that $\sigma(\mathfrak{X})$ means the set of points which is obtained by the rotation σ from $\mathfrak{X}$. With this notation we define

$$\mathfrak{R} = \sum_{\sigma \in A} \sigma(\mathfrak{X}), \quad \mathfrak{S} = \sum_{\sigma \in B} \sigma(\mathfrak{X}), \quad \mathfrak{T} = \sum_{\sigma \in C} \sigma(\mathfrak{X}). \tag{14}$$

We now assert that

1. *The sets $\mathfrak{R}$, $\mathfrak{S}$ and $\mathfrak{T}$ are disjoint.*

2. *Each point of $\mathfrak{U} - \mathfrak{L}$ belongs to one of the three sets.*

We show first that to assume that $\mathfrak{R}$ and $\mathfrak{S}$ have a common point leads to a contradiction. If P were a point in the intersection $\mathfrak{R}\mathfrak{S}$, we would have

$$P = \sigma_1(M) = \sigma_2(N),$$

where

$$\sigma_1 \in A, \quad \sigma_2 \in B, \quad M \in \mathfrak{X}, \quad N \in \mathfrak{X}$$

and $M \neq N$ in any case, since otherwise $\sigma_2^{-1}\sigma_1(M) = M$ would be true and M would have to be in $\mathfrak{L}$. And also from $\sigma_1(M) = \sigma_2(N)$ for $M \neq N$ we get a contradiction. For $N = \sigma_2^{-1}\sigma_1(M)$ would hold and therefore $N = \tau(M)$. But this is impossible since M and N as elements of $\mathfrak{X}$ are representatives of *different* classes. Therefore the intersection of $\mathfrak{R}$ and $\mathfrak{S}$ must be empty. Similarly one shows that

$$\mathfrak{R}\mathfrak{T} = 0, \quad \mathfrak{S}\mathfrak{T} = 0.$$

Now we prove assertion 2. Let P be any point of $\mathfrak{U} - \mathfrak{L}$. Then t belongs in any case to $\mathfrak{H}(P)$. Now, if M is the representative of the class $\mathfrak{H}(P)$ lying in $\mathfrak{X}$, $P \in \mathfrak{H}(M)$ also, i.e. $P = \sigma(M)$ for a certain σ from $\mathfrak{G}$. However, σ must belong to one of the three classes A, B or C and hence from definition (14) P belongs to $\mathfrak{R}$, $\mathfrak{S}$

or $\mathscr{T}$. Further for the sets $\mathscr{R}, \mathscr{S}$ and $\mathscr{T}$ defined by (14)

$$\phi(\mathscr{R}) = \mathscr{S} + \mathscr{T}, \quad \psi(\mathscr{R}) = \mathscr{S}, \quad \psi^2(\mathscr{R}) = \mathscr{T} \tag{15}$$

are now valid.

First of all let us prove the first of these relations. Let $P \in \mathscr{R}$ so $P = \sigma(M)$ with $M \in \mathscr{X}$ and $\sigma \in A$. Then

$$\phi(P) = \phi\sigma(M) \in \mathscr{S} + \mathscr{T},$$

since $\phi\sigma$ belongs to $B+C$ from lemma B. Now, conversely, let $Q \in \mathscr{S} + \mathscr{T}$, therefore

$$Q = \rho(M) \text{ with } M \in \mathscr{X}, \rho \in B+C.$$

Therefore $\phi^{-1}Q = \phi^{-1}\rho(M)$. If we put $\phi^{-1}\rho = \sigma$ or $\rho = \phi\sigma$, by lemma B, since $\phi\sigma \in B+C$, it follows that $\sigma \in A$ and therefore $\sigma(M) \in \mathscr{R}$. And this means that $\phi^{-1}(Q) \in \mathscr{R}$, $Q \in \phi(\mathscr{R})$. Hence, we have in fact

$$\phi(\mathscr{R}) = \mathscr{S} + \mathscr{T}.$$

Similarly one shows, using lemma B, that the other relations of (15) are also correct. The details of these arguments are left to the reader.

Let us summarise:—We have discovered a division of the spherical surface in which the sets $\mathscr{Q}, \mathscr{R}, \mathscr{S}$ and $\mathscr{T}$ are mutually disjoint. The congruences stated in Theorem 38 are obtained from (15).

3. Paradoxical Decompositions in $\mathscr{R}_3$

From Hausdorff's theorem we obtain the paradoxical decomposition of two sets which lie on the sphere and have the same cardinality as the continuum: this is, from Theorem 38,

$$\mathscr{S} + \mathscr{T} \equiv \mathscr{S}, \quad \mathscr{S} + \mathscr{T} \equiv \mathscr{T}, \quad \mathscr{S} \cdot \mathscr{T} = 0.$$

By the definition on p. 141 this decomposition is paradoxical. However, paradoxical decompositions of other simpler sets may also be derived from Theorem 38. To derive these we require two more lemmas.

Lemma C: *Let $\mathscr{Q}$ be any point set which is at most denumerable, on the sphere $\mathscr{K}$. Then*

$$\mathscr{K} \overset{2}{\equiv} \mathscr{K} - \mathscr{Q} \tag{16}$$

is true.

The proof of (16) is carried out similarly to that of (4). Here σ is replaced by $\mathscr{Q}$ and $\sigma(\alpha)$ by $\mathscr{Q}(\alpha)$. One only needs to choose the

angle for the rotations about the z-axis, say, so that the sets

$$\mathcal{Q}, \mathcal{Q}(\alpha), \mathcal{Q}(2\alpha), \ldots$$

are all disjoint. One can easily see that this is always possible. Let us denote by η_ν the plane through the z-axis and any point $P_\nu \in \mathcal{Q}$, and by α_ν the angle which the plane η_ν makes with the xz-plane. Then α may be chosen so that the relation

$$\alpha_\nu + n\alpha = \alpha_{\nu'} + 2\pi\mu \tag{17}$$

is not satisfied for any quadruple (n, ν, ν', μ) of non-negative integral numbers. Since the set of α satisfying (17) is denumerable, one can certainly choose the angle of rotation so that all the sets $\mathcal{Q}(\nu\alpha)$ are disjoint.

Hence one can prove lemma C in the same way as equation (4) above.

Lemma D: *If*

$$\mathcal{A} \supset \mathcal{E} \supset \mathcal{B} \text{ and } \mathcal{A} \overset{n}{\equiv} \mathcal{B}, \tag{18}$$

then

$$\mathcal{A} \overset{n+1}{\equiv} \mathcal{E}.$$

By hypothesis the sets $\mathcal{A}$ and $\mathcal{B}$ can be decomposed as follows:—

$$\mathcal{A} = \mathcal{A}_1 + \mathcal{A}_2 + \ldots + \mathcal{A}_n, \quad \mathcal{B} = \mathcal{B}_1 + \mathcal{B}_2 + \ldots + \mathcal{B}_n, \tag{19}$$

where

$$\mathcal{A}_i \mathcal{A}_k = \mathcal{B}_i \mathcal{B}_k = 0 \text{ for } 1 \leqq i < k \leqq n \tag{20}$$

and

$$\mathcal{A}_k \equiv \mathcal{B}_k \text{ for } k = 1, 2, \ldots, n. \tag{21}$$

Thus there is a congruence transformation $\Phi_k(P)$ for each $k = 1, 2, 3, \ldots, n$ which transforms the set $\mathcal{A}_k$ into $\mathcal{B}_k$. Let us combine these transformations by defining

$$\Phi(P) = \Phi_k(P) \text{ for } P \in \mathcal{A}_k.$$

From (18) we have $\Phi(\mathcal{A}) = \mathcal{B} \subset \mathcal{E} \subset \mathcal{A}$ and therefore

$$\Phi^l(\mathcal{A}) \subset \mathcal{A}, \quad \Phi^l(\mathcal{A} - \mathcal{E}) \subset \mathcal{A} \quad (l \text{ a positive integer}). \tag{22}$$

Now we define a new set $\mathcal{A}'$ by the rule

$$\mathcal{A}' = (\mathcal{A} - \mathcal{E}) + \Phi(\mathcal{A} - \mathcal{E}) + \Phi^2(\mathcal{A} - \mathcal{E}) + \ldots \tag{23}$$

and

$$\mathcal{A}'' = \mathcal{A} - \mathcal{A}', \quad \mathcal{E}' = \Phi(\mathcal{A}'). \tag{24}$$

Then from (22) we have that $\mathscr{A}' \subset \mathscr{A}$ and

$$\mathscr{A} = \mathscr{A}' + \mathscr{A}'', \quad \mathscr{A}'\mathscr{A}'' = 0. \tag{25}$$

Further

$$\mathscr{E}' = \Phi(\mathscr{A}') \subset \Phi(\mathscr{A}) = \mathscr{B} \subset \mathscr{E}, \tag{26}$$

therefore

$$\mathscr{E}' \subset \mathscr{E}. \tag{26'}$$

In addition, from (23), it follows that

$$\Phi(\mathscr{A}') = \Phi(\mathscr{A} - \mathscr{E}) + \Phi^2(\mathscr{A} - \mathscr{E}) + \Phi^3(\mathscr{A} - \mathscr{E}) + \ldots \tag{27}$$

therefore

$$\mathscr{A}' = (\mathscr{A} - \mathscr{E}) + \Phi(\mathscr{A}') = (\mathscr{A} - \mathscr{E}) + \mathscr{E}' = \mathscr{A} - (\mathscr{E} - \mathscr{E}').$$

Hence, from (24),

$$\mathscr{E} - \mathscr{E}' = \mathscr{A} - \mathscr{A}' = \mathscr{A}''$$

or

$$\mathscr{E} = \mathscr{E}' + \mathscr{A}'', \quad \mathscr{E}'\mathscr{A}'' = 0. \tag{28}$$

After these preliminaries we can decompose $\mathscr{A}$ into $(n+1)$ mutually disjoint sets:

$$\mathscr{A} = \mathscr{A}' + \mathscr{A}'' = \mathscr{A}'\mathscr{A} + \mathscr{A}'' = \mathscr{A}'\mathscr{A}_1 + \mathscr{A}'\mathscr{A}_2 + \ldots + \mathscr{A}'\mathscr{A}_n + \mathscr{A}''. \tag{29}$$

Now for the terms $\mathscr{A}'\mathscr{A}_k$ of the sum (29)

$$\mathscr{A}'\mathscr{A}_k \equiv \mathscr{E}'\mathscr{B}_k, \quad k = 1, 2, \ldots, n \tag{30}$$

is true. This can be seen as follows: From the definition of the congruence transformation

$$\Phi_k(\mathscr{A}'\mathscr{A}_k) = \Phi(\mathscr{A}'\mathscr{A}_k) = \Phi(\mathscr{A}') \,.\, \Phi(\mathscr{A}_k) = \mathscr{E}'\mathscr{B}_k.$$

Now we can write for $\mathscr{E}$ a decomposition similar to that for $\mathscr{A}$. From (26) $\mathscr{E}' = \mathscr{E}'\mathscr{B}$ and therefore, from (28) and (19), it follows that

$$\mathscr{E} = \mathscr{E}' + \mathscr{A}'' = \mathscr{E}'\mathscr{B} + \mathscr{A}'' = \mathscr{E}'\mathscr{B}_1 + \mathscr{E}'\mathscr{B}_2 + \ldots + \mathscr{E}'\mathscr{B}_n + \mathscr{A}''. \tag{31}$$

From (28) and (20) the terms of the sum on the right hand side are disjoint. Therefore, from (29), (31) and (30),

$$\mathscr{A} \overset{n+1}{\equiv\!\equiv} \mathscr{E}.$$

We can use the lemma just proved to derive other paradoxes. From Theorem 38 it is possible to obtain a decomposition (5) of the spherical surface. The sets denoted there by $\mathscr{R}$, $\mathscr{S}$ and $\mathscr{T}$ we

can now divide as follows: We have that $\mathfrak{R} \equiv \mathfrak{S} + \mathfrak{T}$. Thus there exists a decomposition $\mathfrak{R} = \mathfrak{R}_1 + \mathfrak{R}_2$ for which

$$\mathfrak{R}_1 . \mathfrak{R}_2 = 0, \quad \mathfrak{R}_1 \equiv \mathfrak{S}, \quad \mathfrak{R}_2 \equiv \mathfrak{T}. \tag{32}$$

On the other hand $\mathfrak{R} \equiv \mathfrak{S}, \mathfrak{R} \equiv \mathfrak{T}$. $\mathfrak{R}_1$ and $\mathfrak{R}_2$ are transformed by the appropriate rotations into the disjoint sets $\mathfrak{S}_1$ and $\mathfrak{S}_2$ or $\mathfrak{T}_1$ and $\mathfrak{T}_2$ respectively:

$$\mathfrak{S} = \mathfrak{S}_1 + \mathfrak{S}_2, \quad \mathfrak{T} = \mathfrak{T}_1 + \mathfrak{T}_2, \quad \mathfrak{S}_1 \mathfrak{S}_2 = \mathfrak{T}_1 \mathfrak{T}_2 = 0, \quad \mathfrak{R}_\nu \equiv \mathfrak{S}_\nu \equiv \mathfrak{T}_\nu,$$

$$\nu = 1, 2.$$

Hence we have from (5) the following decomposition of the spherical surface $\mathfrak{K}$:

$$\mathfrak{K} = \mathfrak{R}_1 + \mathfrak{R}_2 + \mathfrak{S}_1 + \mathfrak{S}_2 + \mathfrak{T}_1 + \mathfrak{T}_2 + \mathfrak{Q}.$$

Now we put

$$\mathfrak{K}_1 = \mathfrak{R}_1 + \mathfrak{S}_1 + \mathfrak{T}_1, \quad \mathfrak{K}_2 = \mathfrak{R}_2 + \mathfrak{S}_2 + \mathfrak{T}_2 + \mathfrak{Q}.$$

But from (5′) and (32) the congruences

$$\mathfrak{R}_1 \equiv \mathfrak{R}_2 \equiv \mathfrak{S} \equiv \mathfrak{T} \equiv \mathfrak{R},$$

$$\mathfrak{S}_1 \equiv \mathfrak{S}_2 \equiv \mathfrak{R}_1 \equiv \mathfrak{S},$$

$$\mathfrak{T}_1 \equiv \mathfrak{T}_2 \equiv \mathfrak{R}_1 \equiv \mathfrak{R}_2 \equiv \mathfrak{T}$$

hold. Hence it follows that

$$\mathfrak{K}_1 \overset{3}{\equiv} \mathfrak{K} - \mathfrak{Q}, \quad \mathfrak{K}_2 \overset{4}{\equiv} \mathfrak{K}.$$

Now from lemma C $\mathfrak{K} - \mathfrak{Q} \overset{2}{\equiv} \mathfrak{K}$ and therefore from lemma A $\mathfrak{K}_1 \overset{6}{\equiv} \mathfrak{K}$. Finally we have the result that a *spherical surface is finitely equivalent to two spherical surfaces (of equal radius)*:

$$\mathfrak{K} = \mathfrak{K}_1 + \mathfrak{K}_2, \quad \mathfrak{K}_1 \mathfrak{K}_2 = 0, \quad \mathfrak{K}_1 \overset{6}{\equiv} \mathfrak{K}, \mathfrak{K}_2 \overset{4}{\equiv} \mathfrak{K}. \tag{33}$$

Now it is no longer difficult to prove Theorem 37 formulated on p. 140. In this theorem we are concerned with solid spheres, not spherical surfaces. First of all let us denote by $\mathfrak{k}'$ the solid sphere *without its centre* and note that the outer set of boundary points is the spherical surface $\mathfrak{K}$ just considered. If we replace the points of the spherical surface by the radial lines (open at the centre M of the sphere) through them, we obtain from (33) a corresponding

paradoxical decomposition for the solid sphere (without its centre):

$$k' = k_1' + k_2', \quad k_1' k_2' = 0, \quad k_1' \overset{6}{\equiv} k', \quad k_2' \overset{4}{\equiv} k'. \tag{34}$$

Now, if we include the centre, we get for the solid sphere with its centre:

$$k = k_1' + k_2' + M, \quad k_1 = k_1' + M \overset{(f)}{\equiv} k, \quad k_2' \overset{(f)}{\equiv} k - M.$$

Hence, $k_2' \overset{(f)}{\equiv} k - P$, where P is a point of the spherical surface. Finally from lemma C, it follows for the solid sphere that

$$k = k_1 + k_2', \quad k_1 \overset{(f)}{\equiv} k, \quad k_2' \overset{(f)}{\equiv} k,$$

i.e. Theorem 37 is proved.

Hence we discover

Theorem 39: *If $\mathcal{M}_1$ and $\mathcal{M}_2$ are two bounded sets of three-dimensional space which contain a solid sphere k, they are finitely equivalent.*

To prove this we show first of all that $\mathcal{M}_1$ is finitely equivalent to k. Let r be the radius of the sphere k and $\{\mathcal{W}\}$ a cubical lattice whose cubes have edges $a = \frac{2}{3}\sqrt{3} \cdot r$. Then each cube is inscribed in a sphere with radius equal to r. Let s cubes $\mathcal{W}_\sigma$ ($\sigma = 1, 2, \ldots, s$) cover the set $\mathcal{M}_1$ completely. We have therefore

$$\mathcal{M}_1 \subset \mathcal{W}_1 + \mathcal{W}_2 + \ldots + \mathcal{W}_s.$$

Now $\mathcal{M}_1$ is decomposed further into subsets m_σ which are contained in $\mathcal{W}_\sigma$:

$$\mathcal{M}_1 = m_1 + m_2 + \ldots + m_s, \; m_\sigma \subset \mathcal{W}_\sigma.$$

Now, from Theorem 37, $k \overset{(f)}{\equiv} k_1 + k_2$, where k_1 and k_2 are two separate spheres of radius r. It follows further that $k_2 \overset{(f)}{\equiv} k_3 + k_4$ therefore $k \overset{(f)}{\equiv} k_1 + k_3 + k_4$ and by mathematical induction one shows that there are s separate solid spheres of radius r for which

$$k \overset{(f)}{\equiv} k_1 + k_2 + \ldots + k_s \tag{35}$$

is true.

From the definition of the cubes $\mathcal{W}_\sigma$ and the spheres k_σ it now follows immediately that $\mathcal{M}_1$ is finitely equivalent to a subset of $k_1 + k_2 + \ldots + k_s$. Then from (35) $\mathcal{M}_1$ is also finitely equivalent to

a subset $\mathscr{H}$ of $\mathscr{k}$. Therefore we have

$$\mathscr{H} \subset \mathscr{k} \subset \mathscr{M}_1, \quad \mathscr{H} \overset{(f)}{\equiv} \mathscr{M}_1.$$

Thus from lemma D it follows that $\mathscr{M}_1 \overset{(f)}{\equiv} \mathscr{k}$. Hence we obtain Theorem 39 immediately. For, if a sphere of radius r is contained in $\mathscr{M}_1$ as well as in $\mathscr{M}_2$, from our results we have that

$$\mathscr{M}_1 \overset{(f)}{\equiv} \mathscr{k}, \quad \mathscr{M}_2 \overset{(f)}{\equiv} \mathscr{k},$$

and therefore, from lemma A, $\mathscr{M}_1 \overset{(f)}{\equiv} \mathscr{M}_2$.

From Theorem 39 it follows, moreover, that every polyhedron is finitely equivalent to every sphere.

4. Paradoxical Decompositions in the Plane

The results of the last section raise the question whether there exist paradoxical decompositions in the plane. There are in fact such decompositions.

We give a simple example:—

Let $\mathscr{E}$ be a set of complex numbers of the form

$$\zeta = a_n e^{in} + a_{n-1} e^{i(n-1)} + \ldots + a_1 e^{i} + a_0. \tag{36}$$

where n and a_ν ($\nu = 0, 1, 2, \ldots, n$) are *non-negative integers*. Let $\mathscr{A}$ be the subset of $\mathscr{E}$ for which $a_0 = 0$ and $\mathscr{B} = \mathscr{E} - \mathscr{A}$ the complementary set. It is easily seen that the numbers of the set $\mathscr{E}$ are just those obtained from the number 0 by repeated application of the operations

$$P(z) = z + 1, \quad D(z) = e^{i} \,.\, z.$$

Now, if one applies the transformation $D(z)$, to the set $\mathscr{E}$, one obtains $\mathscr{A}$. If the numbers (36) are multiplied by e^{i}, one again obtains a number of the form (36) but with a power of e^{i} appearing as a factor in each term of the sum. Conversely, multiplying a number

$$\zeta = a_n e^{in} + a_{n-1} e^{i(n-1)} + \ldots + a_1 e^{i}$$

of $\mathscr{A}$ by e^{-i} leads to a number of $\mathscr{E}$; so the congruence of sets $\mathscr{A}$ and $\mathscr{E}$ for multiplication by e^{i} means rotating the set $\mathscr{E}$ through the angle $\alpha = 1$. Likewise $P(z) = z + 1$ maps $\mathscr{E}$ onto β. Since this is a congruence transformation too, we have

$$\mathscr{E} = \mathscr{A} + \mathscr{B}, \quad \mathscr{A} \,.\, \mathscr{B} = 0, \quad \mathscr{E} \equiv \mathscr{A}, \quad \mathscr{E} \equiv \mathscr{B},$$

and therefore a paradoxical decomposition.

Now the set $\mathscr{E}$ is *denumerable and unbounded* in contrast to the sets decomposed in $\mathscr{R}_3$. It is natural to ask whether there are also paradoxical decompositions of bounded plane sets which have the same cardinality as the continuum. Using our results in $\mathscr{R}_3$ one could try to find a paradoxical decomposition of a circle in a similar way to that of the sphere. But it turns out that this is impossible.

In fact

Theorem 40: *There are no paradoxical decompositions of the unit circle or of its non-empty subsets.*

is true.

To prove this let us suppose that $\mathscr{E}$ is a set of points of the complex plane with $|z| = 1$ for which there exist decompositions of the following kind:

$$\mathscr{E} = \mathscr{A} + \mathscr{B}, \quad \mathscr{A} \,.\, \mathscr{B} = 0;$$

$$\mathscr{E} = \mathscr{E}_1 + \mathscr{E}_2 + \ldots + \mathscr{E}_k = \mathscr{E}'_1 + \mathscr{E}'_2 + \ldots + \mathscr{E}'_l; \tag{37}$$

$$\mathscr{A} = \mathscr{A}_1 + \mathscr{A}_2 + \ldots + \mathscr{A}_k, \quad \mathscr{B} = \mathscr{B}_1 + \mathscr{B}_2 + \ldots + \mathscr{B}_l,$$

$$\mathscr{A}_{\kappa_1}\mathscr{A}_{\kappa_2} = \mathscr{B}_{\lambda_1}\mathscr{B}_{\lambda_2} = \mathscr{E}_i\mathscr{E}_j = \mathscr{E}'_i\mathscr{E}'_j = 0,$$

where

$$\mathscr{E}_\kappa \equiv \mathscr{A}_\kappa, \quad \mathscr{E}'_\lambda \equiv \mathscr{B}_\lambda. \tag{38}$$

The congruence transformations in (38) between $\mathscr{E}_\kappa$ and $\mathscr{A}_\kappa$ or $\mathscr{E}'_\lambda$ and $\mathscr{B}_\lambda$ are rotations or reflections. They can be written as follows:

$$\phi_\kappa(z) = \phi_\kappa(e^{i\vartheta}) = e^{i\alpha_\kappa} \,.\, z = e^{i(\alpha_\kappa + \vartheta)} \text{ or } \phi_\kappa(z) = e^{-i(\alpha_\kappa + \vartheta)} \tag{39}$$

$$\psi_\lambda(z) = \psi_\lambda(e^{i\vartheta}) = e^{i\beta\lambda} \,.\, z = e^{i(\beta_\lambda + \vartheta)} \text{ or } \psi_\lambda(z) = e^{-i(\beta_\lambda + \vartheta)}.$$

Since $\mathscr{E}$, $\mathscr{A}$ and $\mathscr{B}$ are decomposed into mutually disjoint subsets by (37), one can also combine the congruence transformations (39) as follows:

$$\Phi(\vartheta) = \phi_\kappa(z) \text{ for } z \in \mathscr{E}_\kappa, \; \Psi(\vartheta) = \psi_\lambda(z) \text{ for } z \in \mathscr{E}'_\lambda.$$

We can assume, without loss of generality, that $z = 1$ belongs to the set $\mathscr{E}$. Then we have that $\Phi(0) \in \mathscr{A}$, $\Psi(0) \in \mathscr{B}$ and therefore $\Phi(0) \neq \Psi(0)$, since the intersection of $\mathscr{A}\mathscr{B}$ is empty. Since $\Phi(x) \neq \Psi(x)$ and $\Phi(x) \neq \Psi(y)$ for all $x \in \mathscr{E}$ and $y \in \mathscr{E}$, we have also

$$\Phi\Phi(0) \neq \Psi\Phi(0), \quad \Phi\Phi(0) \neq \Psi\Psi(0), \quad \Phi\Phi(0) \neq \Phi\Psi(0),$$

$$\Phi\Psi(0) \neq \Psi\Phi(0), \quad \Phi\Psi(0) \neq \Psi\Psi(0), \quad \Psi\Psi(0) \neq \Psi\Psi(0).$$

Therefore the four complex numbers $\Phi\Phi(0)$, $\Phi\Psi(0)$, $\Psi\Phi(0)$, $\Psi\Psi(0)$ are different from one another. Now we consider numbers of the type

$$\tau_s = \Theta_s\Theta_{s-1}\ldots\Theta_1(0), \tag{40}$$

where Θ_σ ($\sigma = 1, 2, \ldots, s$) stands for Φ or Ψ. Now one shows easily by mathematical induction that any two numbers τ_s are different from one another if different transformations (Φ or Ψ) are substituted for Θ_σ for any σ. Since there are two possible substitutions for each σ, we have altogether 2^s different numbers of the form (40). From (39) we can also write these numbers as follows:

$$\tau_s = e^{i\Sigma_s} = e^{i(c_1+c_2+\ldots+c_s)}, \tag{41}$$

where c_σ stands for any $\pm\alpha_\kappa$ or $\pm\beta_\lambda$. However, this means that the $r = 2k+2l$ numbers

$$\pm\alpha_1,\ \pm\alpha_2,\ \ldots,\ \pm\alpha_k,\ \pm\beta_1,\ \pm\beta_2,\ \ldots,\ \pm\beta_l$$

come into the question for c_σ in (41). Let us denote these numbers by u_ρ ($\rho = 1, 2, \ldots, r$). Then Σ_s in (41) is a sum of s numbers u_ρ:

$$\Sigma_s = u_{\nu_1} + u_{\nu_2} + \ldots + u_{\nu_s}.$$

The number u_ρ may appear n_ρ times as a term of this sum. Obviously n_ρ can take only the values 0, 1, 2, ..., s. Therefore there are $s+1$ possible values for n_ρ. Since we have a total of r numbers u_ρ at our disposal, there must be less than $(s+1)^r$ different sums Σ_s. Therefore there are less than $(s+1)^r$ different numbers τ_s. But we proved before that there exist 2^s different numbers of this kind. Therefore we must have $2^s < (s+1)^r$. But this is wrong for large s (with r fixed) since it is well-known that

$$\lim_{s\to\infty} \frac{(s+1)^r}{2^s} = 0.$$

Hence the assumption that there exists a paradoxical decomposition of a (non-empty) set $\mathfrak{S}$ lying in $|z| = 1$ leads to a contradiction.

Similarly one can show ([XI 3], p. 56) that

There are no paradoxical decompositions of a non-empty linear set.

Hence it is shown that there is no theorem for the circle analogous to Theorem 38. It can also be shown † that *paradoxical decompositions are impossible for all bounded plane sets.*

† LINDENBAUM, *Fund. Math.*, **8**, p. 218.

However, this still leaves completely open the possibility that two plane sets are finitely equivalent even though they are not equivalent by dissection in elementary geometry. Therefore it makes sense to pose the squaring problem in the geometry of sets (problem **(10)**, p. 4). It is still unsolved. A series of other open problems in the geometry of sets has been formulated by SIERPINSKI [XI 2], [XI 3]).

Chapter XII

THE METHODOLOGICAL SIGNIFICANCE OF "UNSOLVABLE" PROBLEMS

WOLFGANG BOLYAI (1775-1856), a friend of C. F. GAUSS had spent many years of his life trying to prove Euclid's parallel postulate.† The existence of an unsolved problem like this in geometry was a serious annoyance to him. In this connection he wrote to his son JOHANN:

> *It is inconceivable that this pressing darkness, this eternal solar eclipse, this defect, this continual flaw in pure truth should be allowed in geometry.*

Nowadays, if this is quoted in a lecture on the foundations of geometry, the audience just smiles calmly. How is it that the Hungarian mathematician's emotion, certainly meant seriously, appears so funny to modern students? Herein lies a change in the way of thinking which it is important to investigate.

For the mathematicians of past centuries, trained in Platonic philosophy, the theorems of geometry were statements about the world of ideas and the presence of an important and still unsolved problem meant that the scholar did not have a completely clear view in the world of absolutely valid truths. But the work of BOLYAI'S son (at approximately the same time as the Russian LOBACHEVSKY) showed that the parallel problem *cannot* be solved because it is possible to have a geometry free of contradictions in which all the

† The 5th postulate (" parallel postulate ") of EUCLID is the following:

> " *If a straight line intersects two straight lines and the sum of the interior angles which it makes with them on the same side is smaller than two right angles, the two lines intersect when they are produced on this side.*"

Translator's note: This is equivalent to the usual form, viz.: There is a unique straight line through a given point parallel to a given straight line in a plane.

Euclidean axioms and postulates are valid up to the parallel postulate: this is replaced by the statement that there exist at least two parallels to a line through a point in a plane. Moreover, the existence of " non-Euclidean geometry "† like this meant that the mathematicians in the 19th century learned to regard the nature of their axioms differently. In Hilbert's formalistic mathematics it is no longer a matter of " truth " (in the metaphysical sense), but of " certainty ",‡—the certainty that the construction of the theory from the chosen axioms can never lead to a contradiction. The fact that the parallel postulate cannot be proved has thus lost its terror: Euclid's 5th postulate is simply independent of the others and the other axioms cannot by any means be adapted to prove it.

PLATO called mathematics the " awakener of knowledge ". Now it appears that this is the case for modern mathematics in a different sense:—In mathematics it is possible to determine the possibilities and boundaries of a scientific method with a precision which is not possible for example in the domain of historical scholarship. This is particularly significant if one draws on the modern investigations on the possibility of proving the consistency of mathematical theories. There are possible subjects here for the teaching of mathematics in colleges and schools which have not yet been fully used.§

The unsolvable construction problems ((**8**) and for certain radii (**9**)) considered in this book have a methodological significance similar to that of the parallel problem. For thousands of years it was believed that squaring the circle and trisecting the angle by ruler and compasses *must* be possible. Today we know that ruler and compasses are just not suitable instruments for solving these problems. A modern geometer‖ compares the existence of the unsolvable problems in geometry with the fact that a wood-saw is unsuitable for shaving. The contemporary researcher is in fact more prosaic than WOLFGANG BOLYAI but it is also clear that there is good reason for this change of style.

† See e.g. [XII 1].

‡ HILBERT in his lecture " Über das Unendliche " (" On the infinite ") which he gave in Munich, 1925. Printed in the " Grundlagen der Geometrie " (Foundations of Geometry ") 7th edition, 1930.

§ On this topic see the treatise [XII 3] which is also reprinted in the 2nd edition of [VI 10].

‖ W. BLASCHKE in [XII 2], p. 7.

For the " unsolvable " problems of Chapter VIII the position is somewhat different from that for the construction problems. Here an unsolvable problem exists only because the problems were formulated awkwardly:—In the formulation of problem (7) it is implied that there exists in the set of domains $\mathfrak{D}$ one with smallest surface area. However, the set of surface areas of domains $\mathfrak{D}$ which is bounded from below, does not necessarily have a minimum. But a number set of this kind always has a lower bound. Let us replace problem (7) by the following:

(7c) *In a domain $\mathfrak{D}$ a line AB of length* 1 *may be moved (by translation and rotation) so that the direction of AB is turned through* 360°. *It is required to find the lower limit l of the surface areas of all the domains $\mathfrak{D}$.*

This problem is solvable. From Chapter VIII we have $l = 0$ and we can add that, in this case, the lower limit is not a minimum.

From the unsolvability of problem (7) one sees that one may never assume the existence of a solution as obvious for extremal problems if it concerns sets containing an infinite number of elements. And this also means that indirect proofs for extremal problems need to be completed by an existence statement. We illustrate this using the " isoperimetric problem " as an example. Here we are concerned with the following classical problem:

(7d) *Let M be the set of plane, simple, closed, continuous curves of length L. Which of the curves of the set M encloses a domain of greatest surface area?*

There is a beautiful and simple proof by JACOB STEINER [V 4] which says that a non-circular curve does not have this extremal property. Each such curve $\mathfrak{C}$ can be " improved " by giving a curve $\mathfrak{C}'$ of equal length L, which includes a domain $\mathfrak{D}'$ whose surface area $A(\mathfrak{D}')$ is greater than that of the domain $\mathfrak{D}$ enclosed by the curve $\mathfrak{C}$. Hence STEINER deduced that the problem (7c) is solved by the circle of radius $\frac{L}{2\pi}$. WEIERSTRASS was the first to show that the conclusion is not correct. Of course it is not stated that the Steiner extremal problem generally has a solution. First of all it has only been shown that, *if* there exists a solution of this problem, it is the circle. Of course, one can complete Steiner's argument in

this case by giving an " existence proof " or show directly in another way that the circle is the solution of the isoperimetric problem.†

We merely emphasise by an example given first by O. PERRON, that Steiner's method of reasoning is inadmissible.

We can " prove " that 1 is the greatest of all natural numbers by the method used for the isoperimetric problem. This is done as follows: By squaring, one associates with each natural number n different from 1 a greater natural number n^2. Each natural number different from 1 can therefore (by giving a greater number n^2) be " improved " just as above in the geometric proof the closed curves were " improved " by giving another with greater area. This is true for all the natural numbers except 1. Therefore 1 is the greatest natural number.

Here the fallacy is obvious: we have implied the existence of a greatest natural number, and this is certainly wrong.

† See in this connection, e.g. [II 7].

BIBLIOGRAPHY

The bibliography is in the order of the chapters of the book. Those works which contain full particulars about the literature on the subject are indicated by a star at the number.

II 1. W. KILLING and H. HOVESTADT. *Handbuch des mathematischen Unterrichts* I. Leipzig and Berlin, 1910.

II 2. L. FEJES TOTH. Abschätzung des kürzesten Abstandes zweier Punkte eines auf der Kugel liegenden Punktsystems, *Jber. dt. MatVerein*, **53**, pp. 66-68, 1943.

II 3. D. HILBERT and S. COHN-VOSSEN. *Anschauliche Geometrie.* New York, 1944.

II 4. W. LIETZMANN. *Elementare Kugelgeometrie mit numerischen und konstruktiven Methoden.* Göttingen, 1949.

II 5. W. HABICHT and B. VAN DER WAERDEN. Lagerungen von Punkten auf der Kugel., *Math. Annln.*, **123**, pp. 223-234, 1951.

II 6. B. VAN DER WAERDEN. Punkte auf der Kugel. Drei Zusätze., *Math. Annln.*, **125**, pp. 213-222, 1952.

II 7.*L. FEJES TOTH. *Lagerungen in der Ebene, auf der Kugel und im Raum.* Berlin, Göttingen, Heidelberg, 1953.

II 8. L. FEJES TOTH. Kreisüberdeckungen in der hyperbolischen Ebene., *Acta Math Ac. Sc. Hung.*, IV, 1-2, 1953.

II 9. L. FEJES TOTH. Kreisausfüllungen in der hyperbolischen Ebene., *Acta Math. Ac. Sc. Hung.*, IV, 1-2, 1953.

II 10. H. MESCHKOWSKI. Elementare Behandlung von Lagerungsproblemen. *Math.-phys. SemesterB.*, **4**, pp. 256-262, 1955.

II 11. A. HEPPES. Über mehrfache Kreislagerungen., *Elem. Math.* **10**, pp. 125-127, 1955.

III 1. K. SCHÜTTE and B. VAN DER WAERDEN. Auf welcher Kugel haben 5, 6, 7 oder 8 Punkte im Minimalabstand 1 Platz?, *Math. Annln.*, **123**, pp. 96-123, 1951.

III 2. K. SCHÜTTE. Überdeckung der Kugel mit 8 Kreisen., *Math. Annln.*, **129**, pp. 101-106, 1955.

IV 1. F. SUPNICK. On the dense packing of spheres, *Trans. Amer. math. Soc.*, **65**, pp. 14-26, 1949.

IV 2. M. E. WISE. Dense random packing of unequal spheres, *Philips Res. Rep.*, **7**, pp. 321-343, 1952.

IV 3. A. H. BOERDIJK. Some remarks concerning close-packing of equal spheres, *Philips Res. Rep.*, **7**, pp. 303-313, 1952.

IV 4. H. HADWIGER. Einlagerung kongruenter Kugeln in eine Kugel., *Elem. Math.*, **7**, pp. 97-103, 1952.

V 1. H. W. E. JUNG. Über die kleinste Kugel, die eine räumliche Figur einschließt., *(Crellcs) J. reine angew. Math.*, **123**, pp. 241-257, 1901.

V 2. H. W. E. JUNG. Über den kleinsten Kreis, der eine ebene Figur einschließt., *(Crelles) J. reine angew. Math.*, **137**, pp. 310-313, 1910.

V 3. J. PÁL. Über ein elementares Variationsproblem, *Kgl. danske Vidensk. Selskab Mat.-fys. Meddr.*, III, 2, 1920.
V 4. H. RADEMACHER and O. TOEPLITZ. *Von Zahlen und Figuren.* Berlin, 1933. (*The Enjoyment of Mathematics.* Princeton, 1957.)
V 5. R. SPRAGUE. Über ein elementares Variationsproblem, *Mat. Tideskrift*, 1936, pp. 96-99.
V 6. D. GALE. On inscribing n-dimensional sets in a regular n-simplex. *Proc. Am. Math. Soc.*, **4**, pp. 222-225, 1953.
V 7. A. KIRSCH. Die Pferchkugel eines Punkthaufens, *Math.-phys. SemesterBer*, **3**, pp. 214-218, 1953.
V 8. H. G. EGGLESTON. Covering the three-dimensional set with sets of smaller diameter, *J. London math. Soc.*, **30**, pp. 11-24, 1955.
V 9. H. HADWIGER and H. DEBRUNNER. Ausgewählte Einzelprobleme der kombinatorischen Geometrie in der Ebene., *Enseign. math.* **1**, pp. 56-89, 1955.
V 10.*J. M. JAGLOM and W. G. BOLTJANSKI. *Konvexe Figuren.* Berlin, 1956.
V 11. B. GRÜNBAUM. A simple proof of Borsuk's conjecture in three dimensions, *Proc. Camb. phil. Soc.*, **53**, pp. 776-778, 1957.

VI 1. M. DEHN. Über den Rauminhalt., *Math. Annln.*, **55**, pp. 465-478, 1901.
VI 2. M. ZACHARIAS. *Elementargeometrie der Ebene und des Raumes.* Berlin, 1930.
VI 3. J. P. SYDLER. Sur la décomposition des polyèdres, *Comment. math. helvet.*, **16**, pp. 266-273, 1943.
VI 4. H. HADWIGER. Bemerkungen zur elementaren Inhaltslehre des Raumes, *Elem. Math.*, **4**, pp. 3-7, 1949.
VI 5. E. KAMKE. *Mengenlehre.* Berlin, 1949.
VI 6. H. HADWIGER. Zerlegungsgleichheit und additive Polyederfunktionale, *Comment. math. hely.*, **24**, pp. 204-218, 1950.
VI 7. J. P. SYDLER. Sur les conditions nécessaires pour l'équivalence des polyèdres euclidiens, *Elem. Math.*, **7**, pp. 49-53, 1952.
VI 8. J. P. SYDLER. Sur l'équivalence des polyèdres à dièdres rationnels, *Elem. Math.*, **8**, pp. 75-79, 1953.
VI 9. H. HADWIGER. Der Inhaltsbegriff, seine Begründung und Wandlung in alterer und neuerer Zeit., *Mitt. naturf. Ges. Bern.*, **11**, pp. 13-41, 1954.
VI 10.*H. MESCHKOWSKI. *Wandlungen des mathematischen Denkens*, Braunschweig, 1956.
VI 11. D. HILBERT. *Grundlagen der Geometrie.* 8. *Aufl.* Stuttgart, 1956.
VI 12.*H. HADWIGER. *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie.* Berlin, Göttingen, Heidelberg, 1957.
VI 13. M. GOLDBERG. Tetrahedra equivalent to cubes by dissection, *Elem. Math.*, **13**, pp. 107-109, 1958.

VII 1. M. DEHN. Über die Zerlegung von Rechtecken in Rechtecke, *Math. Annln.*, **57**, pp. 314-332, 1903.
VII 2. M. KRAITCHIK. *La Mathématique des Jeux.* Brussels, 1930.
VII 3. A. STÖHR. Über Zerlegungen von Rechtecken in inkongruente Quadrate, *Schr. Math. Inst. und Inst. angew. Math. Univ. Berlin*, **4**, pp. 119-140, 1939.
VII 4. R. SPRAGUE. Beispiel einer Zerlegung des Quadrates in lauter verschiedene Quadrate, *Math. Z.*, **45**, pp. 607-608, 1939.
VII 5. R. SPRAGUE. Über die Zerlegung von Rechtecken in lauter verschiedene Quadrate, *J. reine angew. Math.*, **182**, pp. 60-64, 1940.

VII 6. R. SPRAGUE. Zur Abschätzung der Mindestzahl inkongruenter Quadrate, die ein gegebenes Rechteck ausfüllen, *Math. Z.*, **46**, pp. 460-471, 1940.
VII 7. R. L. BROOKS, C. A. B. SMITH, A. H. STONE and W. T. TUTTE. The dissection of rectangles into squares, *Duke J. of Math.*, **7**, pp. 312-340, 1940.
VII 8. C. J. BOUWKAMP. On the dissection of rectangles into squares, *Kon. Ned. Ak. van Wet.*, **49**, pp. 1176-1188, 1946; **50**, pp. 58-78, pp. 1296-1299, 1947.
VII 9. W. T. TUTTE. Squaring the square, *Canad. J. Math.*, pp. 197-209, 1950.
VII 10. M. GOLDBERG. The squaring of developable surfaces, *Scr. math.*, **18**, pp. 17-24, 1952.

VIII 1. E. TROST. Bemerkungen zu einem Satz über Mengen von Punkten mit ganzzahliger Entfernung, *Elem. Math.*, 6, pp. 59-60, 1951.
VIII 2. M. ALTWEGG. Ein Satz über Mengen mit ganzzahliger Entfernung. *Elem. Math.*, **7**, pp. 56-58, 1952.
VIII 3. A. MÜLLER. Auf einem Kreis liegende Punktmengen ganzzahliger Entfernungen, *Elem. Math.*, **8**, pp. 37-38, 1953.
VIII 4. F. STEIGER. Zu einer Frage über Mengen von Punkten mit ganzzahliger Entfernung, *Elem. Math.*, **8**, pp. 66-68, 1953.
VIII 5. H. HADWIGER. Ungelöste Probleme, *Elem. Math.*, **13**, p. 85, 1958.
VIII 6. W. SIERPINSKI. Sur les ensembles de points aux distances rationelles situés sur un cercle, *Elem. Math.*, **14**, pp. 25-27, 1959.

IX 1. O. PERRON. *Algebra*, **1**, **2**. Berlin, 1932.
IX 2. P. BUCHNER. Eine Aufgabe, die mit Zirkel und Lineal nicht lösbar ist, *Elem. Math.*, **2**, pp. 14-16, 1947.
IX 3. E. ROTH-DESMEULES. Noch eine Aufgabe, die mit Zirkel und Lineal nicht lösbar ist. *Elem. Math.*, **3**, pp. 65-67, 1948.
IX 4. B. VAN DER WAERDEN. *Moderne Algebra.* Berlin, Göttingen, Heidelberg, 1950.
IX 5. H. LEBESGUE. *Leçons sur les constructions géométriques.* Paris, 1950.
IX 6. E. KASNER and J. HARRISON. The trisection of horn angles, *Scr. math.*, **17**, pp. 231-235, 1951.
IX 7. L. BIEBERBACH. *Theorie der geometrischen Konstruktionen.* Basle, 1952.

X 1. N. M. NESTOROVIC. On the squaring of the circle and the circularisation of the square in the Lobachevskian plane (Russian), *Dokl. Akad. Nauk. SSSR*, 63, pp. 613-614, 1948.
X 2. C. L. SIEGEL. *Transcendental Numbers.* Princeton, 1949.
X 3. E. BEUTEL. *Die Quadratur des Kreises.* Leipzig, 1951.
X 4. TH. SCHNEIDER. *Einführung in die transzendenten Zahlen.* Berlin, Göttingen, Heidelberg, 1957.

XI 1. W. SIERPINSKI. Sur quelques problèmes concernant la congruence des ensembles de points, *Elem. Math.*, **5**. pp. 1-4, 1950.
XI 2. W. SIERPINSKI. *Algèbre des Ensembles.* Warsaw, 1951.
XI 3. W. SIERPINSKI. On the congruence of sets and their equivalence by finite decomposition, *Lucknow Univ. Studies XX*, Fac. of Science, The Lucknow University, 1954.

XI 4.*J. P. NATANSON. *Theorie der Funktionen einer reellen Veränderlichen.* Berlin, 1954.
XI 5. H. HADWIGER. Ungelöste Probleme Nr. 5, *Elem. Math.*, **10**, p. 66, 1955.
XI 6. P. S. ALEXANDROFF. *Einführung in die Mengenlehre und die Theorie der reellen Funktionen.* Berlin, 1956.

XII 1. H. MESCHKOWSKI. *Nichteuklidische Geometrie.* Brunswick, 1954.
XII 2. W. BLASCHKE. *Reden und Reisen eines Geometers.* Berlin, 1957.
XII 3. H. MESCHKOWSKI. Menschenbildung im Zeitalter der Automatisierung, *Erziehung und Bildung*, **11**, Part 4, 1958.

INDEX

Adjunction, 117

Basis polyhedron, 84 *et seq.*
Besicovitch, 107
Bieberbach, 126 *et seq.*
Blaschke, 160
Bolyai, W. and J., 159
Borsuk, 65
Bouwkamp, 92 *et seq.*
Brooks, 91, 92

Characteristic, 9 *et seq.*
Choquet, 110
Cluster of points, 2, 57 *et seq.*
Compasses, 113 *et seq.*
Concave function, 11
Construction on the sphere, 128 *et seq.*
Construction with ruler and compasses, 113 *et seq.*
Continuum, 89
Convex domain, 57
Convex function, 10, 23
Coverings, 29 *et seq.*
Covering bodies, 67
Cube, 53, 82

D-bodies, 67
Decomposition of rectangles, 91 *et seq.*
  imperfect, 94
  perfect, 94
  simple, 94
Dehn, 73 *et seq.*, 101
Delian problem, 113, 122
Density of packing, 2, 5, 16 *et seq.*, 50 *et seq.*
Diameter, 57
Dual graph, 96

Electrical engineering, interpretation in, 100
Equivalence of polyhedra, 82 *et seq.*
Equivalent by completion, 73
Equivalent by dissection, 2
Euclid, 159
Euler, 9 *et seq.*
Extension fields, 117
Extremal problems, 103 *et seq.*

Face-centred, 54
Fejes Toth, 21
Fields, 115
Free edges, 8 *et seq.*
Functional, 71, 89

Gale, 67, 69
Golden section, 116
Graph, 30 *et seq.*, 95 *et seq.*

Hadwiger, 66, 81, 82, 111, 112
Hamel, 84
Hilbert, 73, 160

Icosahedron, 77
Insertion ruler, 114
Irreducible equation, 125
Irreducible graph, 32, 40

Jensen, 11, 23
Juel, 76, 79
Jung, 2, 67

Kirchhoff, 101
Kraitchik, 91 *et seq.*
Kreweras, 110

Lattice packing, 50 *et seq.*
Lebesgue, 57
Lexell, 18
Lobachevsky, 159

Network, 101
Number field, 115 *et seq.*

Octahedron, 67 *et seq.*, 77
Ohm, 101

Packings, 5 *et seq.*, 29 *et seq.*, 50 *et seq.*
Pál, 57, 60
Parallel postulate, 159
Perron, 162
Plato, 113, 160
Point sets, 57 *et seq.*
Polyhedron, functional of, 88
Polygonal net, 8
Pre-dependent, 85 *et seq.*
Ptolemy, 111

Quinquesection, 117 *et seq.*

Regular packings, 5
Reulaux, 63
Rhombic dodecahedron, 77
Ruler, 113 *et seq.*

Schütte, 1, 38 *et seq.*
Set-theoretical squaring problem, 4
Sierpinski, 140
Simplex, 69
Smith, 91
Spherical compasses, 130 *et seq.*
Sprague, 60 *et seq.*, 91 *et seq.*, 112
Squaring the circle on the sphere, 3
Steiger, 112
Steiner, 161 *et seq.*
Stereographic projection, 128 *et seq.*
Stone, 91
Sydler, 77 *et seq.*

Tetrahedron, 78
Tile, 2, 57 *et seq.*
Triangular polyhedron, 21
Trisection, 121
Tschirnhaus, 122
Tutte, 91

van der Waerden, 1, 38

Width, 62